arXiv:1010.6231v1 [cs.CC] 29 Oct 2010

A POLYNOMIAL-TIME ALGORITHM FOR ESTIMATING THE PARTITION FUNCTION OF THE FERROMAGNETIC ISING MODEL ON A REGULAR MATROID

LESLIE ANN GOLDBERG AND MARK JERRUM

Abstract. We investigate the computational difficulty of approximating the partition function of the ferromagnetic Ising model on a regular matroid. Jerrum and Sinclair have shown that there is a fully polynomial randomised approximation scheme (FPRAS) for the class of graphic matroids. On the other hand, the authors have previously shown, subject to a complexity-theoretic assumption, that there is no FPRAS for the class of binary matroids, which is a proper superset of the class of graphic matroids. In order to map out the region where approximation is feasible, we focus on the class of regular matroids, an important class of matroids which properly includes the class of graphic matroids, and is properly included in the class of binary matroids. Using Seymour's decomposition theorem, we give an FPRAS for the class of regular matroids.

1. Introduction

Classically, the Potts model [11] in statistical physics is defined on a graph. Let q be a positive integer and $G=(V,E)$ a graph with edge weights $\boldsymbol{\gamma}=\{\gamma_e : e\in E\}$; the weight $\gamma_e>-1$ represents a "strength of interaction" along edge e. The q-state Potts partition function specified by this weighted graph is

$$Z_{\text{Potts}}(G;q,\boldsymbol{\gamma})=\sum_{\sigma:V\to[q]}\prod_{e=\{u,v\}\in E}\big(1+\gamma_e\,\delta(\sigma(u),\sigma(v))\big), \tag{1}$$

where $[q]=\{1,\dots,q\}$ is a set of q spins or colours, and $\delta(s,s')$ is 1 if $s=s'$, and 0 otherwise. The partition function is a sum over "configurations" σ which assign spins to vertices in all possible ways. We are concerned with the computational complexity of approximately evaluating the partition function (1) and generalisations of it. For reasons that will become apparent shortly, we shall be concentrating on the case $q=2$, which is the familiar Ising model. In this special case, the two spins correspond to two possible magnetisations at a vertex (or "site"), and the edges (or "bonds") model interactions between sites. In the *ferromagnetic* case, when $\gamma_e>0$, for all $e\in E$, the configurations σ with many adjacent like spins make a greater contribution to the partition function $Z_{\text{Potts}}(G;q,\boldsymbol{\gamma})$ than those with few; in the *antiferromagnetic* case, when $\gamma_e<0$, the opposite is the case.

The work described in this paper was partly supported by EPSRC Research Grant (refs EP/I011528/1 & EP/I011935/1) "Computational Counting".

An equivalent way of looking at (1) is as a restriction of the (multivariate) Tutte polynomial, which is defined as follows:

$$\widetilde{Z}(G; q, \boldsymbol{\gamma}) = \sum_{A \subseteq E} \gamma_A \, q^{\kappa(A) - |V|}, \tag{2}$$

where $\gamma_A = \prod_{e \in A} \gamma_e$ and $\kappa(A)$ denotes the number of connected components in the graph (V, A) [13, (1.2)]. This is perhaps not the most usual expression for the multivariate Tutte polynomial of a graph, but it conveniently generalises to the Tutte polynomial of a matroid (5), also [13, (1.3)], which is the main subject of this article. Although (2) and (1) are formally quite different, they agree when q is a positive integer, up to a factor of $q^{-|V|}$.

Jaeger, Vertigan and Welsh [7] were the first to consider the computational complexity of computing the Tutte polynomial. They considered the classical bivariate Tutte polynomial in which the edge weights are constant, i.e., $\gamma_e = \gamma$ for all $e \in E$. Their approach was to fix q and γ, and consider the computational complexity of computing (2) as a function of the instance graph G. Jaeger et al. showed, amongst other things, that computing the Tutte polynomial exactly is #P-hard when $q > 1$ and $\gamma \in (-1, \infty) - \{0\}$. In particular, this means that the partition function (1) of the Potts model is computationally intractable, unless P = #P.

In the light of this intractability result, it is natural to consider the complexity of approximate computation in the sense of "fully polynomial approximation scheme" or FPRAS. Before stating the known results, we quickly define the relevant concepts. A *randomised approximation scheme* is an algorithm for approximately computing the value of a function $f : \Sigma^* \to \mathbb{R}$. The approximation scheme has a parameter $\varepsilon > 0$ which specifies the error tolerance. A *randomised approximation scheme* for f is a randomised algorithm that takes as input an instance $x \in \Sigma^*$ (e.g., for the problem of computing the bivariate Tutte polynomial of a graph, the graph G) and a rational error tolerance $\varepsilon > 0$, and outputs a rational number z (a random variable of the "coin tosses" made by the algorithm) such that, for every instance x,

$$\Pr\left[e^{-\varepsilon} f(x) \le z \le e^{\varepsilon} f(x)\right] \ge \frac{3}{4}\,. \tag{3}$$

The randomised approximation scheme is said to be a *fully polynomial randomised approximation scheme*, or *FPRAS*, if it runs in time bounded by a polynomial in $|x|$ and ε^{-1}. Note that the quantity 3/4 in Equation (3) could be changed to any value in the open interval $(\frac{1}{2}, 1)$ without changing the set of problems that have randomised approximation schemes [9, Lemma 6.1].

For the problem of computing the bivariate Tutte polynomial in the antiferromagnetic situation, i.e., $-1 < \gamma < 0$, there is no FPRAS for (2) unless RP = NP [4]. This perhaps does not come as a great surprise, since, in the special case $\gamma = -1$, the equivalent expression (1) counts proper colourings of a graph. So

we are led to consider the ferromagnetic case, $\gamma > 0$. Even here, Goldberg and Jerrum [3] have recently provided evidence of computational intractability (under a complexity theoretic assumption that is stronger than RP $\neq$ NP) when $q > 2$.

The sequence of results so far described suggest we should focus on the special case $q = 2$ and $\gamma > 0$, i.e., the ferromagnetic Ising model. Here, at last, there is a positive result to report, as Jerrum and Sinclair [8] have presented an FPRAS for the partition function (2), with $q = 2$ and arbitrary positive weights $\boldsymbol{\gamma}$. As hinted earlier, the Tutte polynomial makes perfect sense in the much wider context of an arbitrary matroid (see Sections 2 and 3 for a quick survey of matroid basics, and of the Tutte polynomial in the context of matroids). The Tutte polynomial of a graph is merely the special case where the matroid is restricted to be graphic. It is natural to ask whether the positive result of [8] extends to a wider class of matroids than graphic. One extension of graphic matroids is to the class of binary matroids. Goldberg and Jerrum [5] recently provided evidence of computational intractability of the ferromagnetic Ising model on binary matroids, under the same strong complexity-theoretic assumption mentioned earlier.

Sandwiched between the graphic (computationally easy) and binary matroids (apparently computationally hard) is the class of regular matroids. Since it is interesting to locate the exact boundary of tractability, we consider here the computational complexity of estimating (in the FPRAS sense) the partition function of the Ising model on a regular matroid. We show that there is an FRPAS in this situation. (See Section 2 for matroid definitions and Section 3 for the definition of the Tutte polynomial $\widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma})$ of a matroid $\mathcal{M}$.)

Theorem 1. *There is an FPRAS for the following problem.*

Instance: *A binary matrix representing a regular matroid* $\mathcal{M}$. *A set* $\boldsymbol{\gamma} = \{\gamma_e : e \in E(\mathcal{M})\}$ *of non-negative rational edge weights.*

Output: $\widetilde{Z}(\mathcal{M}; 2, \boldsymbol{\gamma})$.

Aside from the existing FPRAS for graphic matroids, which also works, by duality, for so-called cographic matroids, the main ingredient in our algorithm is Seymour's decomposition theorem for regular matroids. This theorem has been applied on at least one previous occasion to the design of a polynomial-time algorithm. Golynski and Horton [6] use the approach in their algorithm for finding a minimum-weight basis of the cycle space (or circuit space) of a regular matroid. The decomposition theorem states that every regular matroid is either graphic, cographic, a special matroid on 10 elements named R_{10}, or can be decomposed as a certain kind of sum (called "1-sum", "2-sum" or "3-sum") of two smaller regular matroids (see Theorem 3). Since we know how to handle the base cases (graphic, cographic and R_{10}), it seems likely that the decomposition theorem will yield a polynomial time algorithm quite directly. However, there is a catch (which does not arise in [6]). When we pull apart a regular matroid into two smaller ones, say into the 3-sum of $\mathcal{M}_1$ and $\mathcal{M}_2$, four subproblems are generated for each of the

parts $\mathcal{M}_1$ and $\mathcal{M}_2$. This is fine if the decomposition is fairly balanced at each step, but that is not always the case. In the case of a highly unbalanced decomposition, we face combinatorial explosion.

The solution we adopt is to "solve" recursively only the smaller subproblem, say $\mathcal{M}_2$. Then we construct a constant size matroid $\mathcal{I}_3$ that is equivalent to $\mathcal{M}_2$ in the context of any 3-sum. We then glue $\mathcal{I}_3$ onto $\mathcal{M}_1$, using the 3-sum operation, in place of $\mathcal{M}_2$. The matroid $\mathcal{I}_3$ is small, just six elements, and has the property that forming the 3-sum with $\mathcal{M}_1$ leaves $\mathcal{M}_1$ unchanged as a matroid, though it acquires some new weights from $\mathcal{I}_3$. (In a sense, $\mathcal{I}_3$ is an identity for the 3-sum operation.) Then we just have to find the partition function of $\mathcal{M}_1$ (with amended weights). Since we have four recursive calls on "small" subproblems, but only one on a "large" one, we achieve polynomially bounded running time.

The fact that $\mathcal{I}_3$ is able to simulate the behaviour of an arbitrary regular matroid in the context of a 3-sum is a fortunate accident of the specialisation to $q = 2$. Since the existing algorithm for the graphic case also makes very particular use of the fact that $q = 2$, one gets the impression that the Ising model has very special properties compared with the Potts model in general.

2. Matroid preliminaries

A matroid is a combinatorial structure that has a number of equivalent definitions, but the one in terms of a rank function is the most natural here. A set E (the "ground set") together with a rank function $r : E \to \mathbb{N}$ is said to be a *matroid* if the following conditions are satisfied for all subsets $A, B \subseteq E$: (i) $0 \le r(A) \le |A|$, (ii) $A \subseteq B$ implies $r(A) \le r(B)$ (monotonicity), and (iii) $r(A \cup B) + r(A \cap B) \le r(A) + r(B)$ (submodularity). A subset $A \subseteq E$ satisfying $r(A) = |A|$ is said to be *independent*; a maximal (with respect to inclusion) independent set is a *basis*, and a minimal dependent set is a *circuit*. A circuit with one element is a *loop*. We denote the ground set of matroid $\mathcal{M}$ by $E(\mathcal{M})$ and its rank function by $r_{\mathcal{M}}$. To every matroid $\mathcal{M}$ there is a *dual matroid* $\mathcal{M}^*$ with the same ground set $E = E(\mathcal{M})$ but rank function $r_{\mathcal{M}^*}$ given by $r_{\mathcal{M}^*}(A) = |A| + r_{\mathcal{M}}(E - A) - r_{\mathcal{M}}(E)$. A *cocircuit* in $\mathcal{M}$ is a set that is a circuit in $\mathcal{M}^*$; equivalently, a cocircuit is a minimal set that intersects every basis. A cocircuit with one element is a *coloop*. A thorough exposition of the fundamentals (and beyond) of matroid theory can be found in Oxley's book [10].

We give names to two specific matroids which we will use later. $\mathcal{N}_1$ is defined to be the matroid with ground set $E(\mathcal{N}_1) = \{p\}$ in which $r_{\mathcal{N}_1}(p) = 1$. $\mathcal{N}_3$ is defined to be the matroid with ground set $E(\mathcal{N}_3) = \{p_1, p_2, p_3\}$ in which $E(\mathcal{N}_3)$ is a circuit.

Important operations on matroids include contraction and deletion. Suppose $T \subseteq E$ is any subset of the ground set of matroid $\mathcal{M}$. The *contraction* $\mathcal{M}/T$ *of* T *from* $\mathcal{M}$ is the matroid on ground set $E - T$ with rank function given by $r_{\mathcal{M}/T}(A) = r_{\mathcal{M}}(A \cup T) - r_{\mathcal{M}}(T)$, for all $A \subseteq E - T$. The *deletion* $\mathcal{M}\backslash T$ *of* T *from* $\mathcal{M}$ is the matroid on ground set $E - T$ with rank function given by $r_{\mathcal{M}/T}(A) = r_{\mathcal{M}}(A)$,

for all $A \subseteq E - T$. These operations are often combined, and we write $\mathcal{M}/T\backslash S$ for the matroid obtained by contracting T from $\mathcal{M}$ and then deleting S from the result. The operations of contraction and deletion are dual in the sense that $(\mathcal{M}\backslash T)^* = \mathcal{M}^*/T$. For compactness, we shall often miss out set brackets, writing $\mathcal{M}/p_1\backslash p_2, p_3$, for example, in place of $\mathcal{M}/\{p_1\}\backslash\{p_2, p_3\}$. The *restriction* $\mathcal{M} \mid S$ of $\mathcal{M}$ to $S \subseteq E$ is the matroid on ground set S that inherits its rank function from $\mathcal{M}$; another way of expressing this is to say $\mathcal{M} \mid S = \mathcal{M}\backslash(E - S)$.

The matroid axioms are intended to abstract the notion of linear independence of vectors. Some matroids can be represented concretely as a matrix M with entries from a field K, the columns of the matrix being identified with the elements of E. The rank $r(A)$ of a subset $A \subseteq E$ is then just the rank of the submatrix of M formed from columns picked out by A. A matroid that can be specified in this way is said to be *representable over* K. A matroid that is representable over GF(2) is *binary*, and one that is representable over every field is *regular*; the regular matroids form a proper subclass of binary matroids. Another important class of matroids are ones that arise as the *cycle matroid* of an undirected (multi)graph $G = (V, E)$. Here, the edge set E of the graph forms the ground set of the matroid, and the rank of a subset $A \subseteq E$ is defined to be $r(A) = |V| - \kappa(A)$, where $\kappa(A)$ is the number of connected components in the subgraph (V, A). A matroid is *graphic* if it arises as the cycle matroid of some graph, and cographic if its dual is graphic. Both the class of graphic matroids and the class of cographic matroids are strictly contained in the class of regular matroids.

Now for some definitions more specific to the work in this article. A *cycle* in a matroid is any subset of ground set that can be expressed as a disjoint union of circuits. We let $\mathcal{C}(\mathcal{M})$ denote the set of cycles of a matroid $\mathcal{M}$. If $\mathcal{M}$ is a binary matroid, the symmetric difference of any two cycles is again a cycle [10, Thm. 9.1.2(vi)], so $\mathcal{C}(\mathcal{M})$, viewed as a set of characteristic vectors on $E(\mathcal{M})$, forms a vector space over GF(2), which we refer to as the *circuit space* of $\mathcal{M}$. Indeed, any vector space generated by a set of vectors in $\mathrm{GF}(2)^E$ can be regarded as the circuit space of a binary matroid on ground set E (see the remark following [10, Cor. 9.2.3]). The set of cycles $\mathcal{C}(\mathcal{M})$ of a matroid $\mathcal{M}$ determines the set of circuits of $\mathcal{M}$ (these being just the minimal non-empty cycles), which in turn determines $\mathcal{M}$. Note, from the definition of "cycle", that

$$\mathcal{C}(\mathcal{M}\backslash T) = \{C \in \mathcal{C}(\mathcal{M}) \mid C \subseteq E\backslash T\}.$$

The term "cycle" in this context is not widespread, but is used by Seymour [12] in his work on matroid decomposition; the term "circuit space" is more standard.

Consider two binary matroids $\mathcal{M}_1$ and $\mathcal{M}_2$ with $E(\mathcal{M}_1) = E_1 \cup T$ and $E(\mathcal{M}_2) = E_2 \cup T$ and $E_1 \cap E_2 = \emptyset$. The *delta-sum* of $\mathcal{M}_1$ and $\mathcal{M}_2$ is the matroid $\mathcal{M}_1 \triangle \mathcal{M}_2$ on ground set $E_1 \cup E_2$ with the following circuit space:

$$\mathcal{C}(\mathcal{M}_1 \triangle \mathcal{M}_2) = \big\{C \subseteq E_1 \cup E_2 : C = C_1 \oplus C_2 \text{ for some } C_1 \in \mathcal{C}(\mathcal{M}_1) \text{ and } C_2 \in \mathcal{C}(\mathcal{M}_2)\big\},$$

where $\oplus$ denotes symmetric difference [12]. (The right-hand side of the above equation defines a vector space over GF(2), and hence does describe the circuit space of some matroid.) A situation of particular interest occurs when $\mathcal{M}_1 \mid T = \mathcal{M}_2 \mid T = N$ (say), i.e., $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same restriction, namely N, to T. To investigate this situation, consider the following collection of cycles, this time on the larger set $E_1 \cup T \cup E_2$:

(4)
$$\mathcal{C}(\mathcal{A}) = \big\{C \subseteq E_1 \cup T \cup E_2 : C = C_1 \oplus C_2 \text{ for some } C_1 \in \mathcal{C}(\mathcal{M}_1) \text{ and } C_2 \in \mathcal{C}(\mathcal{M}_2)\big\},$$

For the same reason as before, this is the circuit space of a matroid, which we'll denote by $\mathcal{A}$. The following lemma shows that $\mathcal{A}$ is an *amalgam* of $\mathcal{M}_1$ and $\mathcal{M}_2$; that is to say $\mathcal{A} \mid (E_1 \cup T) = \mathcal{M}_1$ and $\mathcal{A} \mid (E_2 \cup T) = \mathcal{M}_2$. Note that there is no reason a priori to suppose that an amalgam exists, though the lemma shows that it does in this case.

Lemma 2. *Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ are binary matroids with $\mathcal{M}_1 \mid T = \mathcal{M}_2 \mid T$, where $T = E(\mathcal{M}_1) \cap E(\mathcal{M}_2)$. The the matroid $\mathcal{A}$ with circuit space given by (4) (with $E_1 = E(\mathcal{M}_1) - T$ and $E_2 = E(\mathcal{M}_2) - T$) is an amalgam of $\mathcal{M}_1$ and $\mathcal{M}_2$.*

Proof. Certainly, any cycle in $\mathcal{M}_1$ is also a cycle in $\mathcal{A} \mid (E_1 \cup T)$. Now consider any cycle in $\mathcal{A} \mid (E_1 \cup T)$; this cycle has the form $C_1 \oplus C_2$ where $C_1 \in \mathcal{C}(\mathcal{M}_1)$, $C_2 \in \mathcal{C}(\mathcal{M}_2)$ and $C_2 \subseteq T$. But since $\mathcal{M}_1 \mid T = \mathcal{M}_2 \mid T$, we see that C_2 is also a cycle in $\mathcal{M}_1$, and hence $C = C_1 \oplus C_2$ is a cycle in $\mathcal{M}_1$. Thus, $\mathcal{C}(\mathcal{A} \mid (E_1 \cup T)) = \mathcal{C}(\mathcal{M}_1)$, and $\mathcal{A} \mid (E_1 \cup T) = \mathcal{M}_1$. By symmetry, a similar identity holds for $\mathcal{M}_2$, and we deduce that $\mathcal{A}$ is an amalgam of $\mathcal{M}_1$ and $\mathcal{M}_2$. $\square$

There is a routine method that gives an upper bound on the rank function of the amalgam $\mathcal{A}$ in terms of the constituent matroids $\mathcal{M}_1$ and $\mathcal{M}_2$. Since

$$\mathcal{C}(A \backslash T) = \{C \in \mathcal{C}(A) \mid C \subseteq E_1 \cup E_2\} = \mathcal{C}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2),$$

we know $\mathcal{M}_1 \bigtriangleup \mathcal{M}_2 = \mathcal{A} \backslash T$. So the rank function of $\mathcal{A}$ immediately gives us a handle on the rank function of the delta-sum $\mathcal{M}_1 \bigtriangleup \mathcal{M}_2$. We shall make use of this computational tool in the following section.

The special case of the delta-sum when $T = \emptyset$ is called the *1-sum* of $\mathcal{M}_1$ and $\mathcal{M}_2$. (It is just the direct sum of the two matroids.) The special case when T is a singleton that is not a loop or coloop in $\mathcal{M}_1$ or $\mathcal{M}_2$, and $|E(\mathcal{M}_1)|, |E(\mathcal{M}_2)| \geq 3$, is called the *2-sum* of $\mathcal{M}_1$ or $\mathcal{M}_2$ and is denoted $\mathcal{M}_1 \oplus_2 \mathcal{M}_2$. Finally, the special case when $|T| = 3$, T is a circuit in both $\mathcal{M}_1$ and $\mathcal{M}_2$ but contains no co-circuit of either, and $|E(\mathcal{M}_1)|, |E(\mathcal{M}_2)| \geq 7$, is called the *3-sum* of $\mathcal{M}_1$ or $\mathcal{M}_2$ and is denoted $\mathcal{M}_1 \oplus_3 \mathcal{M}_2$ [12].

Our main tool is the following celebrated result of Seymour [12, (14.3)]:

Theorem 3. *Every regular matroid $\mathcal{M}$ may be constructed by means of 1- 2- and 3-sums, starting with matroids each isomorphic to a minor of $\mathcal{M}$, and each either graphic, cographic or isomorphic to a certain 10-element matroid R_{10}.*

It is important for us that a polynomial-time algorithm exists for finding the decomposition promised by this theorem. As Truemper notes [15], such an algorithm is given implicitly in Seymour's paper. However, much more efficient algorithms are known. In particular, [14, (10.6.1)] gives a polynomial-time algorithm for testing whether a binary matroid is graphic or cographic. Also, [15] gives a cubic algorithm for expressing any regular matroid that is not graphic, cographic or isomorphic to R_{10} as a 1-sum, 2-sum or 3-sum. For an exposition of this result, [14, (8.4.1)] gives a polynomial-time algorithm that takes as input a matrix representing a binary matroid and produces a 1-sum or a 2-sum decomposition, or, if the matroid is 3-connnected, produces a 3-sum decomposition if one exists. From the proof of [12, (14.3)], and the conditions on the matroid, (that it be regular, but not graphic, cographic or isomorphic to R_{10}), the matroid has a 3-separation in which the parts are sufficiently large, so by [14, (8.3.12)], the desired 3-sum decomposition does exist (so is constructed).

Truemper's definitions are slightly different from Seymour's (which we use). However, it is easy to check that Truemper's 1-sum [14, Section 8.2] is a 1-sum in the sense of Seymour. It can also be checked that Truemper's 2-sum is a 2-sum in the sense of Seymour. (For this it helps to note that the ground-set element in $E(\mathcal{M}_1) \cap E(\mathcal{M}_2)$ is represented by the column corresponding to element "x" in Truemper's matrix $I \mid B^1$ and by the column corresponding to element "y" in his matrix $I \mid B^2$.) Truemper's 3-sums [14, Section 8.2] are not quite 3-sums in the sense of Seymour, but, as Golynski and Horton have noted [6], it is easy to apply an exchange operation [14, Section 8.5] (in linear time) to obtain a 3-sum in the sense of Seymour.

While Seymour's decomposition theorem is in terms of 1-sums, 2-sums and 3-sums, it will be convenient for us to do our preparatory work, in the following section, using the slightly more general notion of a delta-sum.

3. Tutte polynomial and decomposition

Suppose $\mathcal{M}$ is a matroid, q is an indeterminate, and $\boldsymbol{\gamma} = \{\gamma_e : e \in E(\mathcal{M})\}$ is a collection of indeterminates, indexed by elements of the ground set of $\mathcal{M}$. The (multivariate) *Tutte polynomial* of $\mathcal{M}$ and $\boldsymbol{\gamma}$) is defined to be

$$\widetilde{Z}(\mathcal{M}; q, \boldsymbol{\gamma}) = \sum_{A \subseteq E(\mathcal{M})} \gamma_A \, q^{-r_{\mathcal{M}}(A)}, \tag{5}$$

where $\gamma_A = \prod_{e \in A} \gamma_e$ [13, (1.3)]. In this article, we are interested in the *Ising model*, which corresponds to the specialisation of the above polynomial to $q = 2$, so we shall usually omit the parameter q in the above notation and assume that q is set to 2. As we are concerned with approximate computation, we shall invariably be working in a environment each of the indeterminates γ_e is assigned some real value; furthermore, we shall be focusing on the *ferromagnetic* case, in which those values or *weights* are all non-negative. For convenience, we refer to the pair $(\mathcal{M}, \boldsymbol{\gamma})$ as a

"weighted matroid", and we will assume throughout that $\gamma_e \geq 0$ for all $e \in E(\mathcal{M})$. For background material on the multivariate Tutte polynomial, and its relation to the classical 2-variable Tutte polynomial, refer to Sokal's expository article [13].

In order to exploit Theorem 3, we need to investigate how definition (5) behaves when $\mathcal{M} = \mathcal{M}_1 \triangle \mathcal{M}_2$.

Suppose that E_1, T and E_2 are mutually disjoint sets and consider two binary matroids $\mathcal{M}_1$ and $\mathcal{M}_2$ with $E(\mathcal{M}_1) = E_1 \cup T$ and $E(\mathcal{M}_2) = E_2 \cup T$. Assume, as we did in Section 2, that $\mathcal{M}_1$ and $\mathcal{M}_2$ have the same common restriction $N = \mathcal{M}_1 \mid T = \mathcal{M}_2 \mid T$ to T. Then, as we saw in Section 2, the delta-sum $\mathcal{M} = \mathcal{M}_1 \triangle \mathcal{M}_2$ may be expressed as $\mathcal{M} = \mathcal{A} \backslash T$ where $\mathcal{A}$ is an amalgam of $\mathcal{M}_1$ and $\mathcal{M}_2$.

Oxley [10, Section 12.4] has given an upper bound on the rank function of any amalgam of $\mathcal{M}_1$ and $\mathcal{M}_2$. Thus, the upper bound applies to $r_{\mathcal{A}}$, and hence $r_{\mathcal{M}}$. In particular,

$$\forall X \subseteq E_1 \cup E_2, r_{\mathcal{M}}(X) = r_{\mathcal{A}}(X) \leq \min\{\eta(Y) : Y \supseteq X\}, \tag{6}$$

where

$$\eta(Y) = r_{\mathcal{M}_1}(Y \cap E(\mathcal{M}_1)) + r_{\mathcal{M}_2}(Y \cap E(\mathcal{M}_2)) - r_N(Y \cap T).$$

For $A \subseteq E_i$ and $S \subseteq T$, let $e_i(A, S) = r_{\mathcal{M}_i}(A \cup S) - r_{\mathcal{M}_i}(A)$. Intuitively, this is the "excess" rank that S adds to A. For $A \subseteq E_1$ and $B \subseteq E_2$, let $c(A, B) = r_{\mathcal{M}}(A \cup B) - r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B)$. Intuitively, $c(A, B)$ is the appropriate "correction" to the rank function of $\mathcal{M}$ relative to the rank function of the simple 1-sum of $\mathcal{M}_1$ and $\mathcal{M}_2$.

Equation (6) gives

$$\begin{aligned} r_{\mathcal{M}}(A \cup B) &\leq \min_{S \subseteq T} \left(r_{\mathcal{M}_1}(A \cup S) + r_{\mathcal{M}_2}(B \cup S) - r_N(S)\right) \\ &= r_{\mathcal{M}_1}(A) + r_{\mathcal{M}_2}(B) + \min_{S \subseteq T} \left(e_1(A, S) + e_2(B, S) - r_N(S)\right), \end{aligned}$$

so Equation (6) gives

$$c(A, B) \leq \min_{S \subseteq T} \left(e_1(A, S) + e_2(B, S) - r_N(S)\right). \tag{7}$$

From the definitions of the matroids $\mathcal{N}_1$ and $\mathcal{N}_3$, we can now see that, when $N = \mathcal{N}_1$,

$$c(A, B) \leq \min(0, e_1(A, p) + e_2(B, p) - 1). \tag{8}$$

Also, when $N = \mathcal{N}_3$,

$$c(A, B) \leq \min_{S \subset T} \left(e_1(A, S) + e_2(B, S) - |S|\right). \tag{9}$$

Here we use the fact that the right-hand side of (7) cannot have a *unique* minimum at $S = T$, in order to restrict the minimisation to strict subsets of T.)

The interaction between the classical *bivariate* Tutte polynomial and 2- and 3-sums has been investigated previously, for example by Andrzejak [1]. In principle, it would be possible to assure oneself that his proof carries over from the bivariate

to the multivariate situation, and then recover the appropriate formulas (Lemmas 4 and 5 below) by appropriate algebraic translations. However, in the case of the 3-sum there is an obstacle. While Andrzejak's identities remain valid, as identities of rational functions on q, they become degenerate (through division by zero) under the specialisation $q = 2$. In fact, this degeneracy is crucial to us, in that it reduces the dimension of the bilinear form in Lemma 5 from five, as in Andrzejak's general result, to four, in our special formula for $q = 2$. In light of these considerations, and because certain intermediate results in our proofs are in any case required later, we derive the required formulas here. First the formula for (a slight relaxation of) the 2-sum.

Lemma 4. *Suppose $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ and $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$ are weighted binary matroids with $E(\mathcal{M}_1) \cap E(\mathcal{M}_2) = \{p\}$, where p is not a loop in either $\mathcal{M}_1$ or $\mathcal{M}_2$. Denote by $\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 \bigtriangleup \boldsymbol{\gamma}_2$ the weighting on $E(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2)$ inherited from $\mathcal{M}_1$ and $\mathcal{M}_2$. Then*

$$\widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2; \boldsymbol{\gamma}) = (\widetilde{Z}(\mathcal{M}_1 \backslash p; \boldsymbol{\gamma}_1), \widetilde{Z}(\mathcal{M}_1 / p; \boldsymbol{\gamma}_1)) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \widetilde{Z}(\mathcal{M}_2 \backslash p; \boldsymbol{\gamma}_2) \\ \widetilde{Z}(\mathcal{M}_2 / p; \boldsymbol{\gamma}_2) \end{pmatrix}. \tag{10}$$

Proof. The machinery of the proof is a little heavy in relation to the scale of the result, but its use will provide a warm-up for the proof of the analogous result for 3-sums (Lemma 5).

Write $E(\mathcal{M}_1)$ as $E_1 \cup \{p\}$ and $E(\mathcal{M}_2)$ as $E_2 \cup \{p\}$ and note that

$$N = \mathcal{M}_1 \mid \{p\} = \mathcal{M}_2 \mid \{p\} = \mathcal{N}_1.$$

Use the definition of $c(A, B)$ to write the Tutte polynomial of the delta-sum as

$$\widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2; \boldsymbol{\gamma}) = \sum_{A \subseteq E_1} \sum_{B \subset E_2} \gamma_A \gamma_B q^{-r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B)} q^{-c(A,B)}. \tag{11}$$

We have from (8), that

$$c(A, B) \leq \begin{cases} -1 & \text{iff } e_1(A, \{p\}) = 0 \text{ and } e_2(B, \{p\}) = 0; \\ 0 & \text{otherwise.} \end{cases}$$

We show in Appendix 6.1.1 that this upper bound is also a lower bound, so

$$c(A, B) = \begin{cases} -1 & \text{iff } e_1(A, \{p\}) = 0 \text{ and } e_2(B, \{p\}) = 0; \\ 0 & \text{otherwise.} \end{cases}$$

Thus, letting $\hat{z}$ be the column vector $\hat{z}^i = (\hat{z}^i_0, \hat{z}^i_1)^{\mathrm{T}}$, for $i \in \{1, 2\}$, where

$$\hat{z}^i_0 = \sum_{A : A \subseteq E_i, e_i(A, \{p\}) = 0} \gamma_A q^{-r_{\mathcal{M}_i}(A)} \quad \text{and} \quad \hat{z}^i_1 = \sum_{A : A \subseteq E_i, e_i(A, \{p\}) = 1} \gamma_A q^{-r_{\mathcal{M}_i}(A)},$$

we may express identity (11) as

$$\widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2; \boldsymbol{\gamma}) = (\hat{z}^1)^{\mathrm{T}} C \hat{z}^2, \tag{12}$$

where

$$C = \begin{pmatrix} q & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Having expressed the left-hand side of (10) in terms of $\hat{z}^1$ and $\hat{z}^2$, we now wish to do the same for the right-hand side. For $i \in \{1,2\}$, let

$$z^i = (z^i_0, z^i_1)^{\mathrm{T}} = (\widetilde{Z}(\mathcal{M}_i \backslash p; \boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i / p; \boldsymbol{\gamma}_i))^{\mathrm{T}},$$

and observe that

$$\widetilde{Z}(\mathcal{M}_i \backslash p; \boldsymbol{\gamma}_i) = \sum_{A \subseteq E_i} \gamma_A q^{-r_{\mathcal{M}_i}(A)}$$

and

$$\widetilde{Z}(\mathcal{M}_i / p; \boldsymbol{\gamma}_i) = \sum_{A \subseteq E_i} \gamma_A q^{-r_{\mathcal{M}_i}(A)} q^{1 - e_i(A, \{p\})},$$

so

$$z^i = V \hat{z}^i, \tag{13}$$

where

$$V = \begin{pmatrix} 1 & 1 \\ q & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}.$$

(We are using the fact that $\boldsymbol{\gamma}$ agrees with $\boldsymbol{\gamma}_1$ on E_1 and with $\boldsymbol{\gamma}_2$ on E_2.) Now, $\hat{z}^i_0 \geq 0$ and $\hat{z}^i_1 > 0$ (since the weights γ_e are non-negative and the term $A = \emptyset$ contributes positively to the sum defining $\hat{z}^i_1$). Together with (13) these facts imply

$$z^i_0 \leq z^i_1 < 2 z^i_0, \tag{14}$$

an inequality we shall need later. With

$$D = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix},$$

we can express the right-hand side of (10), using transformation (13), as

$$(z^1)^{\mathrm{T}} D z^2 = (V \hat{z}^1)^{\mathrm{T}} D V \hat{z}^2 = (\hat{z}^1)^{\mathrm{T}} (V^{\mathrm{T}} D V) \hat{z}^2. \tag{15}$$

Comparing (12) and (15), we see that to complete the proof, we just need to verify the matrix equation $C = V^{\mathrm{T}} D V$ with C and V specialised to $q = 2$, and this is easily done. $\square$

The same calculation can be carried through for (a slight relaxation of) the 3-sum. Denote by D the matrix

$$D = \begin{pmatrix} 4 & -1 & -1 & -1 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}.$$

Lemma 5. *Suppose that* $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ *and* $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$ *are weighted binary matroids with* $E(\mathcal{M}_1) \cap E(\mathcal{M}_2) = T = \{p_1, p_2, p_3\}$*, and suppose also that* T *is a circuit in both* $\mathcal{M}_1$ *and* $\mathcal{M}_2$*. For* $i = 1, 2$*, let*

$$z^i = \big(\widetilde{Z}(\mathcal{M}_i \backslash T; \boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_1 \backslash p_2, p_3; \boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_2 \backslash p_1, p_3; \boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_3 \backslash p_1, p_2; \boldsymbol{\gamma}_i)\big)^{\mathrm{T}}.$$

Denote by $\boldsymbol{\gamma} = \boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2$ *the weighting on* $E(\mathcal{M}_1 \triangle \mathcal{M}_2)$ *inherited from* $\mathcal{M}_1$ *and* $\mathcal{M}_2$*. Then*

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}) = (z^1)^{\mathrm{T}} D z^2. \tag{16}$$

Proof. Note that $N = \mathcal{M}_1 \mid T = \mathcal{M}_2 \mid T = \mathcal{N}_3$. Use the definition of $c(A, B)$ to write the Tutte polynomial of the delta-sum as

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}) = \sum_{A \subseteq E_1} \sum_{B \subset E_2} \gamma_A \gamma_B q^{-r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B)} q^{-c(A,B)}. \tag{17}$$

Now, to construct an equation analogous to (10), we will work with the minors $\mathcal{M}_i/S_1 \backslash S_2$ for partitions (S_1, S_2) of T. This gives eight different minors of $\mathcal{M}_i$, but some of them are equivalent. In particular, since T is a circuit of $\mathcal{M}_i$ (which means that p_1, p_2 and p_3 are linearly dependent in the binary matrix representing $\mathcal{M}_i$ but any pair of them is independent), any partition (S_1, S_2) with $|S_1| \geq 2$ gives a matroid $\mathcal{M}_i/S_1 \backslash S_2$ which is equivalent to $\mathcal{M}_i/T$. To see this, note that for $X \subseteq E_i$, $r_{\mathcal{M}_i/S_1 \backslash S_2}(X) = r_{\mathcal{M}_i}(X \cup S_1) - r_{\mathcal{M}_i}(S_1) = r_{\mathcal{M}_i}(X \cup T) - 2 = r_{\mathcal{M}_i}(X \cup T) - r_{\mathcal{M}_i}(T)$. Thus, we will use the five minors with $|S_1| \neq 2$.

Let's collect the formulas for the Tutte polynomials of the minors that we'll need.

$$\widetilde{Z}(\mathcal{M}_i \backslash T; \boldsymbol{\gamma}_i) = \sum_{A \subseteq E_i} \gamma_A q^{-r_{\mathcal{M}_i}(A)} \tag{18}$$

$$\widetilde{Z}(\mathcal{M}_i/p_j \backslash T - p_j; \boldsymbol{\gamma}_i) = \sum_{A \subseteq E_i} \gamma_A q^{-r_{\mathcal{M}_i}(A)} q^{1 - e_i(A, \{p_j\})}, \quad \text{for } j \in \{1, 2, 3\},$$

and

$$\widetilde{Z}(\mathcal{M}_i/T; \boldsymbol{\gamma}_i) = \sum_{A \subseteq E_i} \gamma_A q^{-r_{\mathcal{M}_i}(A)} q^{2 - e_i(A, T)}$$

With an eye to the proof of Lemma 4, we next need to understand the function $e_i(A, S)$. It turns out that, with A fixed, $e_i(A, S)$ is completely determined by its value on the on the singleton sets $S = \{p_1\}, \{p_2\}, \{p_3\}$. Not only that, there are just five possible for the values of $e_i(A, S)$ on those singletons. For $i = 1, 2$, the following collection of predicates $\{\varphi_1^i, \ldots, \varphi_4^i\}$ on $A \subseteq E_i$, captures those five possibilities.

P0. $\varphi_0^i(A)$ iff $e_i(A, \{p_1\}) = e_i(A, \{p_2\}) = e_i(A, \{p_3\}) = 0$ (in which case, $e_i(A, S) = 0$ for all S).

P1. $\varphi_1^i(A)$ iff $e_i(A, \{p_1\}) = 0$ and $e_i(A, \{p_2\}) = e_i(A, \{p_3\}) = 1$ (in which case, $e_i(A, S) = 1$ when $|S| \geq 2$).

P2. $\varphi_2^i(A)$ iff $e_i(A,\{p_2\})=0$ and $e_i(A,\{p_1\})=e_i(A,\{p_3\})=1$ (in which case, $e_i(A,S)=1$ when $|S|\geq 2$).
P3. $\varphi_3^i(A)$ iff $e_i(A,\{p_3\})=0$ and $e_i(A,\{p_1\})=e_i(A,\{p_2\})=1$ (in which case, $e_i(A,S)=1$ when $|S|\geq 2$).
P4. $\varphi_4^i(A)$ iff $e_i(A,\{p_1\})=e_i(A,\{p_2\})=e_i(A,\{p_3\})=1$ (in which case, $e_i(A,S)=2$ when $|S|\geq 2$).

Of course, in all cases, $e_i(A,\emptyset)=0$.

First observe that it is not possible that exactly one of $e_i(A,\{p_1\})$, $e_i(A,\{p_2\})$ and $e_i(A,\{p_3\})$ to take the value 1, since this would imply that two members of T are dependent on A, and hence the third would be. Thus, P0–P4 are exhaustive as well as mutually exclusive.

We still need to verify the additional information provided in parentheses. In case P0, $e_i(A,S)=0$ for all S, since all of the elements in T are dependent on A. In case P1 (P2 and P3 are symmetrical) p_1 is is dependent on A, and so $e_i(A,\{p_1,p_2\})=e_i(A,\{p_2\})=1$; then $e_i(A,T)=e_i(A,\{p_1,p_2\})=1$, since p_3 depends on p_1 and p_2. The final parenthetical claim (in P4) is slightly trickier, and relies on the fact that the matroids we are working with are binary. (There is a simple counterexample based on the uniform matroid $U_{4,2}$ for the claim in general.) So consider the elements of A as columns of the representing matrix (vectors over GF(2)), and similarly consider the elements of S as columns. We must rule out the possibility that $e_i(A,S)=1$ for some 2-element subset S, say $S=\{p_1,p_2\}$. So suppose $e_i(A,\{p_1\})=e_i(A,\{p_2\})=e_i(A,\{p_1,p_2\})=1$. Then p_2 is dependent on $A\cup\{p_1\}$, i.e., so (also viewing p_1 and p_2 as vectors over GF(2)), $p_2=p_1+\sum_{e\in B}e$, for some $B\subseteq A$. (Note that p_1 must be included in this expression, otherwise p_2 would be dependent on A.) Since T is a circuit we know that $p_3=p_1+p_2$. Now write $p_3=p_1+p_2=p_1+p_1+\sum_{e\in B}e=\sum_{e\in B}e$. Thus p_3 is dependent on A, and $e_i(A,p_3)=0$, contrary to what we assumed.

By analogy with the proof of Lemma 4 define $\hat{z}^i=(\hat{z}_0^i,\dots,\hat{z}_4^i)^{\mathrm{T}}$ for $i=1,2$ by

$$\hat{z}_k^i=\sum_{A:A\subseteq E_i,\varphi_k^i(A)}\gamma_A q^{-r_{\mathcal{M}_i}(A)}$$

and

$$\begin{aligned}z^i=\big(&\widetilde{Z}(\mathcal{M}_i\backslash T;\boldsymbol{\gamma}_i),\widetilde{Z}(\mathcal{M}_i/p_1\backslash p_2,p_3;\boldsymbol{\gamma}_i),\\&\widetilde{Z}(\mathcal{M}_i/p_2\backslash p_1,p_3;\boldsymbol{\gamma}_i),\widetilde{Z}(\mathcal{M}_i/p_3\backslash p_1,p_2;\boldsymbol{\gamma}_i),\widetilde{Z}(\mathcal{M}_i/T;\boldsymbol{\gamma}_i)\big)^{\mathrm{T}}.\end{aligned}$$

Note that the dimension here is one greater (five instead of four) than in the statement of the lemma, but this extra dimension will disappear towards the end of the proof. As before, we can re-express the left-hand side of (16) in terms of $\hat{z}^1,\hat{z}^2$.

Our first step will be to show that the value $c(A,B)$ depends on the predicates $\varphi_j^1(A)$ and $\varphi_k^2(B)$ but does not depend otherwise on A and B. Thus, we will be

able to define a 5×5 matrix C with $C_{j,k} = q^{-c(A,B)}$, for $0 \le j,k \le 4$, where $A \subseteq E_1$ and $B \subseteq E_2$ are any sets satisfying $\varphi^1_j(A)$ and $\varphi^2_k(B)$. Using Equation (9), we can construct an element-wise lower bound for the matrix C. In particular,

$$C \ge \begin{pmatrix} q^2 & q & q & q & 1 \\ q & q & 1 & 1 & 1 \\ q & 1 & q & 1 & 1 \\ q & 1 & 1 & q & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

For example, if $\varphi^1_1(A)$ and $\varphi^2_2(B)$ are true, then $e_1(A,\{p_1\}) = 0$ and $e_1(A,\{p_2\}) = e_1(A,\{p_3\}) = 1$ so $e_1(A,S) = 1$ when $|S| \ge 2$. Also, $e_2(B,\{p_2\}) = 0$ and $e_2(B,\{p_1\}) = e_2(B,\{p_3\}) = 1$ so $e_2(B,S) = 1$ when $|S| \ge 2$. Now Equation (9) says $c(A,B) \le \min_{S\subset T}\left(e_1(A,S) + e_2(B,S) - |S|\right)$. Checking the possibilities for S, we find $c(A,B) \le 0$, so $q^{-c(A,B)} \ge 1$. Thus, the entry in row 1 (the second row from the top) and column 2 (the third column from the left) above is 1. The rest of the entries are filled in in the same way.

We show in Appendix 6.1.2 that this lower bound is also an upper bound, so

$$C = \begin{pmatrix} q^2 & q & q & q & 1 \\ q & q & 1 & 1 & 1 \\ q & 1 & q & 1 & 1 \\ q & 1 & 1 & q & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 & 2 & 2 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Then the left-hand side of (16) may be expressed as

$$\widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2; \boldsymbol{\gamma}) = (\hat{z}^1)^{\mathrm{T}} C \hat{z}^2. \tag{19}$$

Also as before, there is a linear relationship between z^i and $\hat{z}^i$, expressed by

$$z^i = V\hat{z}^i, \tag{20}$$

where

$$V = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ q & q & 1 & 1 & 1 \\ q & 1 & q & 1 & 1 \\ q & 1 & 1 & q & 1 \\ q^2 & q & q & q & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 1 & 1 \\ 2 & 1 & 1 & 2 & 1 \\ 4 & 2 & 2 & 2 & 1 \end{pmatrix}.$$

It is straightforward to verify correctness of V row by row. For example, the third row expresses the easily checked fact that

$$\widetilde{Z}(\mathcal{M}_i/p_2\backslash p_1, p_3; \boldsymbol{\gamma}_i) = q\hat{z}^i_0 + \hat{z}^i_1 + q\hat{z}^i_2 + \hat{z}^i_3 + \hat{z}^i_4.$$

Identity (20) allows us to rewrite the right-hand side of (16) in terms of $\hat{z}^1$ and $\hat{z}^2$. Let D' be the 5×5 matrix obtained from D in the statement of the theorem by

padding it out to the right and below by a single column and row of zeros. Then the right-hand side of (16) may be written as

$$(z^1)^{\mathrm{T}} D z^2 = (V\hat{z}^1)^{\mathrm{T}} D' V \hat{z}^2 = (\hat{z}^1)^{\mathrm{T}} (V^{\mathrm{T}} D' V) \hat{z}^2. \tag{21}$$

Comparing (19) and (21), we see that to complete the proof, we just need to verify the matrix equation

$$C = V^{\mathrm{T}} D' V, \tag{22}$$

with C and V specialised to $q = 2$, and this is easily done. □

Remark 1. *With q regarded as an indeterminate, Equation (22) has a unique solution in which D' is a matrix of full rank, whose entries are rational functions of q. When q is specialised to 2, the rank of matrix C drops to 4, allowing a solution in which D' also has rank 4. When $q = 2$, there is flexibility in the choice of the matrix D', which we exploit in order to reduce the dimension by 1. This is the reason that the right-hand side of (16) is of dimension 4, whereas one would expect dimension 5 a priori. This apparent accident is crucial to the design of the algorithm.*

4. Signatures

The goal of this section is to show that, for every weighted binary weighted matroid $(\mathcal{M}, \boldsymbol{\gamma})$ with distinguished element p, there is a small (2-element) weighted matroid $(\mathcal{I}_2, \boldsymbol{\delta})$ that is equivalent to $\mathcal{M}$ in the following sense: if we replace $(\mathcal{M}, \boldsymbol{\gamma})$ by $(\mathcal{I}_2, \boldsymbol{\delta})$ in the context of any 2-sum, the Tutte polynomial (specialised to $q = 2$) of the 2-sum is changed by a factor that is independent of the context. Moreover, the weighted matroid $(\mathcal{I}_2, \boldsymbol{\delta})$ can readily be computed given a "signature" of $\mathcal{M}$. There is also a 6-element weighted matroid $(\mathcal{I}_3, \boldsymbol{\delta})$ that does a similar job for 3-sums.

To make this precise, let $\mathcal{I}_2$ be the matroid with a 2-element ground set $\{p, e\}$ that form a 2-circuit. $\mathcal{I}_2$ can be viewed as the cycle matroid of a graph consisting of two parallel edges, p and e.

Lemma 6. *Suppose $(\mathcal{M}, \boldsymbol{\gamma})$ is a weighted binary matroid with a distinguished element p that is not a loop. Then there is a weight $d \geq 0$ such that, for every weight function $\boldsymbol{\delta}$ with $\delta_e = d$,*

$$\frac{\widetilde{Z}(\mathcal{I}_2/p; \boldsymbol{\delta})}{\widetilde{Z}(\mathcal{I}_2\backslash p; \boldsymbol{\delta})} = \frac{\widetilde{Z}(\mathcal{M}/p; \boldsymbol{\gamma})}{\widetilde{Z}(\mathcal{M}\backslash p; \boldsymbol{\gamma})}.$$

The value d can be computed from $\widetilde{Z}(\mathcal{M}\backslash p; \boldsymbol{\gamma})$ and $\widetilde{Z}(\mathcal{M}/p; \boldsymbol{\gamma})$ — it does not otherwise depend upon $\mathcal{M}$ or $\boldsymbol{\gamma}$.

Proof. Let $z_0 = \widetilde{Z}(\mathcal{M}\backslash p;\boldsymbol{\gamma})$ and $z_1 = \widetilde{Z}(\mathcal{M}/p;\boldsymbol{\gamma})$, and set $d = 2(z_1 - z_0)/(2z_0 - z_1)$. Inequality (14) implies that the numerator is non-negative, and the denominator is strictly positive. Then

$$\widetilde{Z}(\mathcal{I}_2/p;\boldsymbol{\delta})/\widetilde{Z}(\mathcal{I}_2\backslash p;\boldsymbol{\delta}) = (1+d)/(1+\tfrac{1}{2}d) = z_1/z_0 = \widetilde{Z}(\mathcal{M}/p;\boldsymbol{\gamma})/\widetilde{Z}(\mathcal{M}\backslash p;\boldsymbol{\gamma}).$$

□

Corollary 7. *Suppose $(\mathcal{M}_1,\boldsymbol{\gamma}_1)$ and $(\mathcal{M}_2,\boldsymbol{\gamma}_2)$ are weighted binary matroids with $E(\mathcal{M}_1)\cap E(\mathcal{M}_2) = \{p\}$, where p is not a loop in either $\mathcal{M}_1$ or $\mathcal{M}_2$. Then there is a weight $d \geq 0$ such that, for every weight function $\boldsymbol{\delta}$ with $\delta_e = d$,*

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) = \zeta \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}),$$

where $\zeta = 2(2+d)^{-1}\widetilde{Z}(\mathcal{M}_2\backslash p;\boldsymbol{\gamma}_2)$. The value d can be computed from $\widetilde{Z}(\mathcal{M}_2\backslash p;\boldsymbol{\gamma}_2)$ and $\widetilde{Z}(\mathcal{M}_2/p;\boldsymbol{\gamma}_2)$ — it does not otherwise depend upon $(\mathcal{M}_1,\boldsymbol{\gamma}_1)$ or $(\mathcal{M}_2,\boldsymbol{\gamma}_2)$.

Proof. Let $s = \widetilde{Z}(\mathcal{M}_2/p;\boldsymbol{\gamma}_2)/\widetilde{Z}(\mathcal{M}_2\backslash p;\boldsymbol{\gamma}_2)$. By Lemma 4,

$$\widetilde{Z}(\mathcal{M}_1\triangle\mathcal{M}_2;\boldsymbol{\gamma}_1\triangle\boldsymbol{\gamma}_2) = (\widetilde{Z}(\mathcal{M}_1\backslash p;\boldsymbol{\gamma}_1),\widetilde{Z}(\mathcal{M}_1/p;\boldsymbol{\gamma}_1))\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ s \end{pmatrix}\widetilde{Z}(\mathcal{M}_2\backslash p;\boldsymbol{\gamma}_2).$$

Similarly, using Lemma 6,

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = (\widetilde{Z}(\mathcal{M}_1\backslash p;\boldsymbol{\gamma}_1),\widetilde{Z}(\mathcal{M}_1/p;\boldsymbol{\gamma}_1))\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ s \end{pmatrix}\widetilde{Z}(\mathcal{I}_2\backslash p;\boldsymbol{\delta}).$$

Thus, $\zeta = \widetilde{Z}(\mathcal{M}_2\backslash p;\boldsymbol{\gamma}_2)/\widetilde{Z}(\mathcal{I}_2\backslash p;\boldsymbol{\delta})$. □

It is crucial in Corollary 7 that ζ does not depend at all on $(\mathcal{M}_1,\boldsymbol{\gamma}_1)$. As we shall see presently, an analogous result holds for 3-sums, though the calculations are more involved. For a weighted binary matroid $(\mathcal{M},\boldsymbol{\gamma})$ with distinguished elements $T = \{p_1,p_2,p_3\}$, the *signature* of $(\mathcal{M},\boldsymbol{\gamma})$ with respect to T is the vector

$$\sigma(\mathcal{M};T,\boldsymbol{\gamma}) = \frac{\widetilde{Z}(\mathcal{M}/p_1\backslash p_2,p_3;\boldsymbol{\gamma}),\widetilde{Z}(\mathcal{M}/p_2\backslash p_1,p_3;\boldsymbol{\gamma}),\widetilde{Z}(\mathcal{M}/p_3\backslash p_1,p_2;\boldsymbol{\gamma})}{\widetilde{Z}(\mathcal{M}\backslash T;\boldsymbol{\gamma})}.$$

What we are seeking in the 3-sum case is a small matroid whose signature is equal to that of a given binary matroid $\mathcal{M}$. Such a matroid will be equivalent to $\mathcal{M}$ in the context of any 3-sum. Before constructing such a matroid we give two technical lemmas that investigate inequalities between the Tutte polynomial of various minors of a binary matroid. These inequalities will restrict the domain of possible signatures that can occur.

Lemma 8. *Suppose that $(\mathcal{M},\boldsymbol{\gamma})$ is a weighted binary matroid with distinguished elements $T=\{p_1,p_2,p_3\}$, and that T is a circuit in $\mathcal{M}$. Let*

$$\begin{aligned}z^{\mathrm{T}} &= (z_0,z_1,z_2,z_3,z_4)\\ &= \big(\widetilde{Z}(\mathcal{M}\backslash T;\boldsymbol{\gamma}),\widetilde{Z}(\mathcal{M}/p_1\backslash p_2,p_3;\boldsymbol{\gamma}),\\ &\qquad \widetilde{Z}(\mathcal{M}/p_2\backslash p_1,p_3;\boldsymbol{\gamma}),\widetilde{Z}(\mathcal{M}/p_3\backslash p_1,p_2;\boldsymbol{\gamma}),\widetilde{Z}(\mathcal{M}/T;\boldsymbol{\gamma})\big).\end{aligned}$$

Then the following (in)equalities hold: (i) $z_0>0$, (ii) $z_0\le z_1,z_2,z_3\le 2z_0$, (iii) $\frac12 z_4<z_1,z_2,z_3\le z_4$ and (iv) $z_1+z_2+z_3=2z_0+z_4$.

Proof. Inequality (i) follows from the definition (5) since the contribution of $A=\emptyset$ is 1 and the contribution of every other A is non-negative. Consider identity (20). Since $\hat z\ge 0$ (coordinatewise), the vector z in the statement of the lemma is in the cone generated by the columns of V (i.e., z is a non-negative linear combination of columns). Denote the rows of V by $V_0,\dots,V_4$. Then, interpreting all vector inequalities coordinatewise, (ii) is a consequence of $V_0\le V_1,V_2,V_3\le 2V_0$ (iii), with non-strict inequality, of $\frac12 V_4\le V_1,V_2,V_3\le V_4$, and (iv) of $V_1+V_2+V_3=2V_0+V_4$. We know that $\hat z_4>0$, since the weights are non-negative and its expansion contains at least one non-zero term, namely the one corresponding to $A=\emptyset$; this implies that the first inequality in (iii) must be strict. □

Lemma 9. *Under the same assumptions as Lemma 8, $z_1z_2,z_1z_3,z_2z_3\le z_0z_4$.*

Proof. Let the ground set of $\mathcal{M}$ be $E\cup T$ where where $E\cap T=\emptyset$. Recall the earlier definition

$$\hat z_k=\sum_{A:A\in E,\varphi_k(A)}\gamma_A q^{-r_{\mathcal{M}}(A)},\quad \text{for } 0\le k\le 4,$$

and recall also that the the predicates φ_k are exhaustive and mutually exclusive, so that $\widetilde{Z}(\mathcal{M}\backslash T;\boldsymbol{\gamma})=\hat z_0+\hat z_1+\cdots+\hat z_4$. For brevity, write $\hat z=\widetilde{Z}(\mathcal{M}\backslash T;\boldsymbol{\gamma})$.

The rank function $r_{\mathcal{M}}$ of a matroid is submodular, so $-r_{\mathcal{M}}$ is supermodular and the probability distribution μ (defined by $\mu(A)=q^{-r_{\mathcal{M}}(A)}\gamma(A)/\widetilde{Z}(\mathcal{M}\backslash T,\boldsymbol{\gamma})$) associated with the random cluster model satisfies the condition for Fortuin-Kasteleyn-Ginibre (FKG) inequality [2, Section 19, Theorem 5], provided $q\ge1$. Let $f=1-e(A,\{p_1\})$ and $g=1-e(A,\{p_2\})$. Note that these are both monotonically increasing, as A grows. (This can either be seen directly, or as a consequence of submodularity of $r_{\mathcal{M}}$.) So the quantities f and g are positively correlated, i.e.,

$$\Pr_\mu(fg=1)=\mathrm{E}_\mu(fg)\ge\mathrm{E}_\mu(f)\,\mathrm{E}_\mu(g)=\Pr_\mu(f=1)\Pr_\mu(g=1),$$

which is equivalent to

$$\frac{\hat z_0}{\hat z}\ge\frac{\hat z_0+\hat z_1}{\hat z}\times\frac{\hat z_0+\hat z_2}{\hat z},$$

which in turn is equivalent to

$$\hat z_0\hat z\ge(\hat z_0+\hat z_1)(\hat z_0+\hat z_2). \tag{23}$$

Consider the first inequality we are required to establish, namely $z_0z_4 \geq z_1z_2$. Using the linear transformation (20), we may express this as an equivalent inequality in terms of $\hat{z}_k$:

$$\hat{z}(4\hat{z}_0 + 2(\hat{z}_1 + \hat{z}_2 + \hat{z}_3) + \hat{z}_4) \geq (2\hat{z}_0 + 2\hat{z}_1 + \hat{z}_2 + \hat{z}_3 + \hat{z}_4)(2\hat{z}_0 + \hat{z}_1 + 2\hat{z}_2 + \hat{z}_3 + \hat{z}_4),$$

where the four linear factors may be read off from the appropriate rows (first, last, second and third, respectively) of the matrix V. Applying the definition of $\hat{z}$ we obtain the equivalent inequality

$$\hat{z}(\hat{z} + 3\hat{z}_0 + \hat{z}_1 + \hat{z}_2 + \hat{z}_3) \geq (\hat{z} + \hat{z}_0 + \hat{z}_1)(\hat{z} + \hat{z}_0 + \hat{z}_2),$$

which further simplifies, through cancellation, to

$$\hat{z}_0\hat{z} + \hat{z}_3\hat{z} \geq (\hat{z}_0 + \hat{z}_1)(\hat{z}_0 + \hat{z}_2). \tag{24}$$

Now (23) implies (24), and we are done, since the other two advertised inequalities follow by symmetry. □

Denote by $\mathcal{I}_3$ the 6-element matroid with ground set $\{p_1, p_2, p_3, e_1, e_2, e_3\}$, whose circuit space is generated by the circuits $\{p_1, e_1\}$, $\{p_2, e_2\}$, $\{p_3, e_3\}$, and $\{p_1, p_2, p_3\}$. The matroid $\mathcal{I}_3$ can be thought of as the cycle matroid of a certain graph, namely, the graph with parallel pairs of edges $\{p_1, e_1\}$, $\{p_2, e_2\}$ and $\{p_3, e_3\}$ in which edges p_1, p_2 and p_3 form a length-3 cycle in the graph.

Let $T = \{p_1, p_2, p_3\}$. We start by showing that, as long as a signature (s_1, s_2, s_3) satisfies certain equations, which Lemmas 8 and 9 will guarantee, then it is straightforward to compute a weighting $\boldsymbol{\delta}$ so that the weighted matroid $(\mathcal{I}_3, \boldsymbol{\delta})$ has signature $\sigma(\mathcal{I}_3; T, \boldsymbol{\gamma}) = (s_1, s_2, s_3)$.

Lemma 10. *Suppose s_1, s_2 and s_3 satisfy*

$$2 + s_1 - s_2 - s_3 > 0, \quad 2 - s_1 + s_2 - s_3 > 0, \quad 2 - s_1 - s_2 + s_3 > 0, \tag{25}$$

$$s_1 + s_2 + s_3 - 3 \geq 0, \tag{26}$$

(27)

$$s_1+s_2+s_3-s_2s_3-2 \geq 0, \quad s_1+s_2+s_3-s_1s_3-2 \geq 0, \ \textit{and} \quad s_1+s_2+s_3-s_1s_2-2 \geq 0;$$

then there are non-negative weights d_1, d_2 and d_3 such that, for any weight function $\boldsymbol{\delta}$ with $\delta_{e_1} = d_1$, $\delta_{e_2} = d_2$ and $\delta_{e_3} = d_3$, $\sigma(\mathcal{I}_3; T, \boldsymbol{\delta}) = (s_1, s_2, s_3)$. The values d_1, d_2 and d_3 can be computed from s_1, s_2 and s_3.

Proof. Define

$$\begin{aligned} S_1 &= 2 + s_1 - s_2 - s_3 \\ S_2 &= 2 - s_1 + s_2 - s_3 \\ S_3 &= 2 - s_1 - s_2 + s_3 \\ R &= s_1 + s_2 + s_3 - 2. \end{aligned}$$

Define the weights d_1, d_2, d_3 for $\boldsymbol{\delta}$ as follows:

$$\begin{aligned} d_1 &= -1 + \sqrt{RS_1/S_2S_3} \\ d_2 &= -1 + \sqrt{RS_2/S_1S_3} \\ d_3 &= -1 + \sqrt{RS_3/S_1S_2}, \end{aligned}$$

By inequalities (25) and (26), R, S_1, S_2 and S_3 are all strictly positive. Thus d_1, d_2, d_3 are well defined. Finally

$$RS_1 - S_2S_3 = 4(s_1 + s_2 + s_3 - s_2s_3 - 2) \geq 0,$$

where the inequality is (27). Thus d_1, and hence, by symmetry, d_2 and d_3, are all non-negative.

Let $Y = d_1d_2 + d_1d_3 + d_2d_3 + d_1d_2d_3$, and note that

$$\begin{aligned} \widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta}) &= q^{-0} + q^{-1}(d_1 + d_2 + d_3) + q^{-2}Y, \text{ and} \\ \widetilde{Z}(\mathcal{I}_3/p_1 \backslash p_2, p_3; \boldsymbol{\delta}) &= q^{-0}(1 + d_1) + q^{-1}(d_2 + d_3 + Y). \end{aligned}$$

Substituting for d_1, d_2, d_3 in these expressions, using the helpful identity

$$Y + d_1 + d_2 + d_3 + 1 = (1 + d_1)(1 + d_2)(1 + d_3) = R\sqrt{R/S_1S_2S_3},$$

we obtain

$$\widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta}) = \tfrac{1}{4}(R + S_1 + S_2 + S_3)\sqrt{R/S_1S_2S_3} = \sqrt{R/S_1S_2S_3},$$

and

$$\widetilde{Z}(\mathcal{I}_3/p_1 \backslash p_2, p_3; \boldsymbol{\delta}) = \tfrac{1}{2}(R + S_1)\sqrt{R/S_1S_2S_3} = s_1\sqrt{R/S_1S_2S_3},$$

and hence $\widetilde{Z}(\mathcal{I}_3/p_1 \backslash p_2, p_3; \boldsymbol{\delta})/\widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta}) = s_1$. By symmetry, similar identities hold for the other components of the signature of $(\mathcal{I}_3, \boldsymbol{\delta})$. Summarising, $\sigma(\mathcal{I}_3; T, \boldsymbol{\delta}) = (s_1, s_2, s_3)$, as desired. □

Temporarily leaving aside the issue of approximation, the way that we will use Lemma 10 is captured in the following corollary, which is analogous to Corollary 7.

Corollary 11. *Suppose that $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ and $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$ are weighted binary matroids with $E(\mathcal{M}_1) \cap E(\mathcal{M}_2) = T = \{p_1, p_2, p_3\}$, and suppose also that T is a circuit in both $\mathcal{M}_1$ and $\mathcal{M}_2$. Let $(s_1, s_2, s_3) = \sigma(\mathcal{M}_2; T, \boldsymbol{\gamma}_2)$. Then there are non-negative weights d_1, d_2 and d_3 such that, for any weight function $\boldsymbol{\delta}$ with $\delta_{e_1} = d_1$, $\delta_{e_2} = d_2$ and $\delta_{e_3} = d_3$,*

$$\widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{M}_2; \boldsymbol{\gamma}_1 \bigtriangleup \boldsymbol{\gamma}_2) = \zeta \widetilde{Z}(\mathcal{M}_1 \bigtriangleup \mathcal{I}_3; \boldsymbol{\gamma}_1 \bigtriangleup \boldsymbol{\delta}),$$

where

$$\zeta = \widetilde{Z}(\mathcal{M}_2 \backslash T; \boldsymbol{\gamma}_2)/\widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{M}_2 \backslash T; \boldsymbol{\gamma}_2)/\sqrt{R/S_1S_2S_3},$$

in the notation of the proof of Lemma 10. The values d_1, d_2 and d_3 can be computed from s_1, s_2 and s_3 — they do not otherwise depend upon $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ or $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$. Moreover, the values R, S_1, S_2 and S_3 are byproducts of this computation.

Proof. As before, for $i = 1, 2$, let

$$z^i = \big(\widetilde{Z}(\mathcal{M}_i\backslash T;\boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_1\backslash p_2, p_3;\boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_2\backslash p_1, p_3;\boldsymbol{\gamma}_i), \widetilde{Z}(\mathcal{M}_i/p_3\backslash p_1, p_2;\boldsymbol{\gamma}_i)\big)^{\mathrm{T}}.$$

Let $s = (1, s_1, s_2, s_3)^{\mathrm{T}}$. By Lemma 5,

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) = (z^1)^{\mathrm{T}} D z^2 = \widetilde{Z}(\mathcal{M}_2\backslash T;\boldsymbol{\gamma}_2)(z^1)^{\mathrm{T}} D s.$$

Similarly, if $\sigma(\mathcal{I}_3; T, \boldsymbol{\delta}) = (s_1, s_2, s_3)$, then

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{I}_3\backslash T;\boldsymbol{\delta})(z^1)^{\mathrm{T}} D s.$$

To simplify the notation (and avoid confusing the index "2" in z^2 with an exponent), let z denote the vector z^2. Now, from inequalities (iii) and (iv) of Lemma 8, we see that $2(z_0 + z_1) > z_1 + z_2 + z_3$, so Equation (25) is satisfied. By inequality (ii) of Lemma 8, Equation (26) is satisfied. Finally, from Lemma 9 and identity (iv) of Lemma 8, $z_2 z_3 \le z_0(z_1 + z_2 + z_3 - 2z_0)$, so Equation (27) is satisfied. So, by Lemma 10, the weights d_1, d_2 and d_3 can be computed and $\sigma(\mathcal{I}_3; T, \boldsymbol{\delta}) = (s_1, s_2, s_3)$. The value of $\widetilde{Z}(\mathcal{I}_3\backslash T;\boldsymbol{\delta})$ is calculated in the proof of Lemma 10. □

The problem with using Corollary 11 to replace the complicated expression $\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2)$ with the simpler $\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta})$ is that, in general, we will not be able to compute the necessary values s_1, s_2 and s_3. Instead, we will use our FPRAS recursively to approximate these values. Thus, we need a version of Corollary 11 that accommodates some approximation error. Unfortunately, this creates some technical complexities. We will use the following lemma.

Lemma 12. *Suppose that $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ and $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$ are weighted binary matroids with $E(\mathcal{M}_1) \cap E(\mathcal{M}_2) = T = \{p_1, p_2, p_3\}$, and suppose also that T is a circuit in both $\mathcal{M}_1$ and $\mathcal{M}_2$.*

Suppose that $\varepsilon \le 1$ and that ϱ is a sufficiently small positive constant ($\varrho = 1/6000$ will do). Suppose that, coordinatewise,

$$e^{-\varepsilon\varrho}\sigma(\mathcal{M}_2; T, \boldsymbol{\gamma}_2) \le (\tilde{s}_1, \tilde{s}_2, \tilde{s}_3) \le e^{\varepsilon\varrho}\sigma(\mathcal{M}_2; T, \boldsymbol{\gamma}_2). \tag{28}$$

Then there are non-negative weights d_1, d_2 and d_3 such that, for any weight function $\boldsymbol{\delta}$ with $\delta_{e_1} = d_1$, $\delta_{e_2} = d_2$ and $\delta_{e_3} = d_3$,

$$e^{-\varepsilon}\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) \le \zeta\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) \le e^{\varepsilon}\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2;\boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2),$$

where

$$\zeta = \widetilde{Z}(\mathcal{M}_2\backslash T;\boldsymbol{\gamma}_2)/\sqrt{R/S_1S_2S_3},$$

in the notation of the proof of Lemma 10. The values d_1, d_2 and d_3 can be computed from $\tilde{s}_1$, $\tilde{s}_2$ and $\tilde{s}_3$ — they do not otherwise depend upon $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ or $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$. Moreover, the values R, S_1, S_2 and S_3 are byproducts of this computation.

Proof. Let $(r_1, r_2, r_3) = \sigma(\mathcal{M}_2; T, \boldsymbol{\gamma}_2)$. First, using Lemma 16 in the appendix with $\chi = \varepsilon\varrho$, we can use $\tilde{s}_1$, $\tilde{s}_2$ and $\tilde{s}_3$ to compute s_1, s_2 and s_3 satisfying

$$e^{-66\varepsilon\varrho}s_i \le r_i \le e^{66\varepsilon\varrho}s_i, \tag{29}$$

and Equations (25), (26) and (27). To see that Lemma 16 applies, note that Equation (49) follows from Lemma 8 (ii). Equation (47) is analogous to Equation (25) and is established in the same way as Equation (25) is established in the proof of Corollary 11. Similarly, Equation (48) is analogous to Equation (27) and is established as in the proof of Corollary 11. (It is not clear that $r_1 \le r_2 \le r_3$, but this can be thought of as a renaming inside Lemma 16. Furthermore, the corresponding assumption $\tilde{s}_1 \le \tilde{s}_2 \le \tilde{s}_3$ is without loss of generality, since, if the assumption does not hold, then these values can be swapped without violating the proximity of $\tilde{s}_i$ to r_i.)

By Lemma 10, s_1, s_2 and s_3 can be used to compute non-negative weights d_1, d_2 and d_3 such that, for any weight function $\boldsymbol{\delta}$ with $\delta_{e_1} = d_1$, $\delta_{e_2} = d_2$ and $\delta_{e_3} = d_3$, $\sigma(\mathcal{I}_3; T, \boldsymbol{\delta}) = (s_1, s_2, s_3)$. As in the proof of Corollary 11, this implies

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta})(z^1)^{\mathrm{T}} D s,$$

where $s = (1, s_1, s_2, s_3)^{\mathrm{T}}$ and $\widetilde{Z}(\mathcal{I}_3 \backslash T; \boldsymbol{\delta}) = \sqrt{R/S_1 S_2 S_3}$ for some byproducts R, S_1, S_2 and S_3 of the computation.

Now let $r = (1, r_1, r_2, r_3)^{\mathrm{T}}$. Then, as in the proof of Corollary 11,

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) = \widetilde{Z}(\mathcal{M}_2 \backslash T; \boldsymbol{\gamma}_2)(z^1)^{\mathrm{T}} D r.$$

Also, since $66\varrho \le 1/56$, $e^{-\varepsilon/56} s_i \le r_i \le e^{\varepsilon/56} s_i$. The result then follows from Lemma 15 in the appendix, since z, s and r have positive entries, and satisfy $1 \le z_i/z_0 \le 2$, $1 \le s_i/s_0 \le 2$, and $1 \le r_i/r_0 \le 2$ for $i \in \{1, 2, 3\}$ by Lemma 8 (i) and (ii). □

We also need a lemma, similar to Lemma 12, that is appropriate for 2-sum.

Lemma 13. *Suppose $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ and $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$ are weighted binary matroids with $E(\mathcal{M}_1) \cap E(\mathcal{M}_2) = \{p\}$, where p is not a loop in either $\mathcal{M}_1$ or $\mathcal{M}_2$. Let $z_0 = \widetilde{Z}(\mathcal{M}_2 \backslash p; \boldsymbol{\gamma}_2)$ and let $z_1 = \widetilde{Z}(\mathcal{M}_2 / p; \boldsymbol{\gamma}_2)$. Suppose that $\varepsilon \le 1$ and that ϱ is a sufficiently small positive constant (as in Lemma 12 — here it suffices to take $\varrho \le 1/160$). Suppose that $e^{-\varepsilon\varrho} z_i \le \tilde{z}_i \le e^{\varepsilon\varrho} z_i$ for $i \in \{0, 1\}$. Then there is a weight $d \ge 0$ such that, for every weight function $\boldsymbol{\delta}$ with $\delta_e = d$,*

$$e^{-\varepsilon} \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) \le \zeta \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) \le e^{\varepsilon} \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2),$$

where $\zeta = 2(2+d)^{-1} z_0$. The value d can be computed from $\tilde{z}_0$ and $\tilde{z}_1$. It does not otherwise depend upon $(\mathcal{M}_1, \boldsymbol{\gamma}_1)$ or $(\mathcal{M}_2, \boldsymbol{\gamma}_2)$.

Proof. This is similar to the proof of Lemma 12, but easier. First, observe that $z_0 > 0$ and $1 \le z_1/z_0 \le 2$ (Equation (14)). The first step is to use $\tilde{z}_0$ and $\tilde{z}_1$ to compute z'_0 and z'_1 such that $z'_0 > 0$ and $1 \le z'_1/z'_0 \le 2$ and $e^{-4\varepsilon\varrho} z'_i \le z_i \le e^{4\varepsilon\varrho} z'_i$. This is straightforward. Just set $z'_0 = \tilde{z}_0 e^{\varepsilon\varrho}$. If $\tilde{z}_1 \le z'_0$ then set $z'_1 = z'_0$. If $\tilde{z}_1 \ge 2z'_0$ then set $z'_1 = 2z'_0$. Otherwise, set $z'_1 = \tilde{z}_1$. (Note that $\frac{\tilde{z}_1}{\tilde{z}_0} \ge e^{-2\varepsilon\varrho} \frac{z_1}{z_0} \ge e^{-2\varepsilon\varrho}$, so if $\tilde{z}_1 \le z'_0$ then $z'_1 = e^{\varepsilon\varrho} \tilde{z}_0 \le e^{\varepsilon\varrho} e^{2\varepsilon\varrho} \tilde{z}_1 \le e^{4\varepsilon\varrho} z_1$. Similarly, $\frac{\tilde{z}_1}{\tilde{z}_0} \le e^{2\varepsilon\varrho} \frac{z_1}{z_0} \le 2e^{2\varepsilon\varrho}$, so if $\tilde{z}_1 \ge 2z'_0$ then $z'_1 = 2e^{\varepsilon\varrho} \tilde{z}_0 \ge e^{\varepsilon\varrho} e^{-2\varepsilon\varrho} \tilde{z}_1 \ge e^{-2\varepsilon\varrho} z_1$.) As in the proof of Lemma 6, let

$d = 2(z_1' - z_0')/(2z_0' - z_1')$. Let $s' = z_1'/z_0'$ and note that, for every weight function $\boldsymbol{\delta}$ with $\delta_e = d$, $\frac{\widetilde{Z}(\mathcal{I}_2/p;\boldsymbol{\delta})}{\widetilde{Z}(\mathcal{I}_2\backslash p;\boldsymbol{\delta})} = s'$. Let $s = z_1/z_0$. As in the proof of Corollary 7, note that

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{M}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\gamma}_2) = \big(\widetilde{Z}(\mathcal{M}_1\backslash p; \boldsymbol{\gamma}_1), \widetilde{Z}(\mathcal{M}_1/p; \boldsymbol{\gamma}_1)\big) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ s \end{pmatrix} z_0,$$

and

$$\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = \big(\widetilde{Z}(\mathcal{M}_1\backslash p; \boldsymbol{\gamma}_1), \widetilde{Z}(\mathcal{M}_1/p; \boldsymbol{\gamma}_1)\big) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ s' \end{pmatrix} (1 + \tfrac{d}{2}).$$

Since $e^{-8\varepsilon\varrho} s' \le s \le e^{8\varepsilon\varrho} s'$, similar to the proof of Lemma 15,

$$\begin{aligned} e^{-160\varepsilon\varrho}\big(\widetilde{Z}(\mathcal{M}_1\backslash p; \boldsymbol{\gamma}_1),&\widetilde{Z}(\mathcal{M}_1/p; \boldsymbol{\gamma}_1)\big) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ s \end{pmatrix} \\ &\le \big(\widetilde{Z}(\mathcal{M}_1\backslash p; \boldsymbol{\gamma}_1), \widetilde{Z}(\mathcal{M}_1/p; \boldsymbol{\gamma}_1)\big) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ s' \end{pmatrix} \\ &\le e^{160\varepsilon\varrho}\big(\widetilde{Z}(\mathcal{M}_1\backslash p; \boldsymbol{\gamma}_1), \widetilde{Z}(\mathcal{M}_1/p; \boldsymbol{\gamma}_1)\big) \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ s \end{pmatrix}. \end{aligned}$$

□

4.1. **Simple Delta-sums.** We complete this section by investigating the (simple) way in which the special matroids $\mathcal{I}_2$ and $\mathcal{I}_3$ interact with delta-sums with $|T| = 1$ and $|T| = 3$, respectively

Suppose $(\mathcal{M}, \boldsymbol{\gamma})$ is a weighted binary matroid on a ground set $E(\mathcal{M})$ of m elements, and with distinguished element p, which is not a loop. Consider the delta-sum $\mathcal{M} \triangle \mathcal{I}_2$. The ground set of this matroid also has m elements, and it shares all but one element with $E(\mathcal{M})$. Thus, we have a natural correspondence between the ground sets of the two matroids. We claim that under this correspondence the two matroids are the same, and for this it is enough to verify that they have the same circuit space. Note that $\{p, e\}$ is the unique non-empty cycle in $\mathcal{I}_2$. For any cycle C in $\mathcal{M}$, exactly one of $C' = C$ or $C' = C \oplus \{p, e\}$ is a cycle in $\mathcal{M} \triangle \mathcal{I}_2$. The mapping $C \mapsto C'$ is invertible, and is the required bijection between cycles in $\mathcal{M}$ and those in $\mathcal{M} \triangle \mathcal{I}_2$. Now suppose $\delta_e = d$, and and let $\boldsymbol{\gamma}'$ be derived from $\boldsymbol{\gamma}$ by assigning $\gamma_p = d$. Then $\widetilde{Z}(\mathcal{M} \triangle \mathcal{I}_2, \boldsymbol{\gamma} \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma}')$.

A similar observation applies to $\mathcal{I}_3$ under delta-sum with $|T| = 3$. Suppose $(\mathcal{M}, \boldsymbol{\gamma})$ is a weighted binary matroid with distinguished elements $T = \{p_1, p_2, p_3\}$, where T is a cycle in $\mathcal{M}$. Consider the delta-sum $\mathcal{M} \triangle \mathcal{I}_3$. As before, there is a natural correspondence between the ground sets of the two matroids. Let C_1, C_2, C_3 be the three 2-circuits in $\mathcal{I}_3$ including elements p_1, p_2, p_3, respectively. Any cycle C in $\mathcal{M}$ can be transformed in a unique way to a cycle C' in $\mathcal{M} \triangle \mathcal{I}_3$, by adding a subset of circuits from $\{C_1, C_2, C_3\}$. The mapping $C \mapsto C'$ is invertible, and is a bijection between cycles in $\mathcal{M}$ and those in $\mathcal{M} \triangle \mathcal{I}_3$. Now let $\boldsymbol{\gamma}'$ be derived

Step 1: If $\mathcal{M}$ is graphic, cographic or R_{10} then estimate $\widetilde{Z}(\mathcal{M},\boldsymbol{\gamma})$ directly.

Step 2: Otherwise use Seymour's decomposition algorithm to express $\mathcal{M}$ as $\mathcal{M}_1 \triangle \mathcal{M}_2$, where $\triangle$ is a 1-, 2- or 3-sum. Recall that $E(\mathcal{M}) = E(\mathcal{M}_1) \oplus E(\mathcal{M}_2)$. Let $T = E(\mathcal{M}_1) \cap E(\mathcal{M}_2)$, $E_1 = E(\mathcal{M}_1) - T$ and $E_2 = E(\mathcal{M}_2) - T$. Noting $E(\mathcal{M}) = E_1 \cup E_2$, let $\boldsymbol{\gamma}_1 : E_1 \to \mathbb{R}^+$ and $\boldsymbol{\gamma}_2 : E_2 \to \mathbb{R}^+$ be the restrictions of $\boldsymbol{\gamma}$ to E_1 and E_2. Assume without loss of generality that $|E(\mathcal{M}_2)| \le |E(\mathcal{M}_1)|$ (otherwise swap their names).

Step 3: If $\triangle$ is s 3-sum then let $T = \{p_1, p_2, p_3\}$ (say). Execute Steps 4a–7a in Figure 2. If $\triangle$ is a 2-sum then let $T = \{p\}$. Execute Steps 4b–7b in Figure 3. If $\triangle$ is a 1-sum then recursively estimate $\widetilde{Z}(\mathcal{M}_1; \boldsymbol{\gamma}_1)$ with accuracy parameter $\varepsilon|\mathcal{M}_1|/|\mathcal{M}|$, and $\widetilde{Z}(\mathcal{M}_2; \boldsymbol{\gamma}_2)$ with accuracy parameter $\varepsilon|\mathcal{M}_2|/|\mathcal{M}|$, and return the product of the two, which is an estimate of $\widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma})$.

FIGURE 1. Algorithm for estimating the Ising partition function of a regular matroid $\mathcal{M}$ given accuracy parameter $\varepsilon \le 1$. ϱ is a sufficiently small positive constant which does not depend upon $\mathcal{M}$ or ε. See Lemmas 12 and 13.

from $\boldsymbol{\gamma}$ by assigning $\gamma_{p_1} = \delta_{e_1}$, $\gamma_{p_2} = \delta_{e_2}$ and $\gamma_{p_3} = \delta_{e_3}$. Then $\widetilde{Z}(\mathcal{M} \triangle \mathcal{I}_3, \boldsymbol{\gamma} \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma}')$.

5. The algorithm

We now have all the ingredients for the algorithm for estimating $\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}) = \widetilde{Z}(\mathcal{M};2,\boldsymbol{\gamma})$, given a weighted regular matroid $(\mathcal{M},\boldsymbol{\gamma})$ and an accuracy parameter ε. The base cases for this recursive algorithm are when $\mathcal{M}$ is graphic, cographic or R_{10}. In these cases we estimate $\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma})$ "directly", which means the following. If $\mathcal{M}$ is R_{10} then we evaluate $\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma})$ by brute force. If $\mathcal{M}$ is graphic, we form the weighted graph $(G,\boldsymbol{\gamma})$ whose (weighted) cycle matroid is $(\mathcal{M},\boldsymbol{\gamma})$. Then the partition function of the Ising model on $(G,\boldsymbol{\gamma})$ may be estimated using the algorithm of Jerrum and Sinclair [8]. If $\mathcal{M}$ is cographic, then its dual $\mathcal{M}^*$ is graphic, and

$$\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}) = \gamma_E q^{-r_{\mathcal{M}}(E)} \widetilde{Z}(\mathcal{M}^*;\boldsymbol{\gamma}^*),$$

where $E = E(\mathcal{M})$, and $\boldsymbol{\gamma}^*$ is the dual weighting given by $\gamma^*_e = q/\gamma_e = 2/\gamma_e$ for all $e \in E(\mathcal{M})$ [13, 4.14a]. (Ground set elements e with $\gamma_e = 0$ do not cause any problems, because they can just be deleted.) Then we proceed as before, but using $(\mathcal{M}^*,\boldsymbol{\gamma}^*)$ in place of $(\mathcal{M},\boldsymbol{\gamma})$. The proposed algorithm is presented as Figure 1.

Using the guarantees from Lemmas 12 and 13, it is easy to see that the algorithm is correct. That is, given a regular matroid $\mathcal{M}$ and an accuracy parameter $\varepsilon < 1$, the algorithm returns an estimate $\hat{Z}$ satisfying $e^{-\varepsilon}\widetilde{Z}(\mathcal{M},\boldsymbol{\gamma}) \le \hat{Z} \le e^{\varepsilon}\widetilde{Z}(\mathcal{M},\boldsymbol{\gamma})$.

Step 4a: Recusively estimate $z_0 = \widetilde{Z}(\mathcal{M}_2 \backslash T; \boldsymbol{\gamma}_2)$, $z_1 = \widetilde{Z}(\mathcal{M}_2/p_1 \backslash p_2, p_3; \boldsymbol{\gamma}_2)$, $z_2 = \widetilde{Z}(\mathcal{M}_2/p_2 \backslash p_1, p_3; \boldsymbol{\gamma}_2)$ and $z_3 = \widetilde{Z}(\mathcal{M}_2/p_3 \backslash p_1, p_2; \boldsymbol{\gamma}_2)$ with accuracy parameter $\varepsilon\varrho|\mathcal{M}_2|/(4|\mathcal{M}|)$.

Step 5a: Using Lemma 12, compute d_1, d_2 and d_3 such that, for any weight function $\boldsymbol{\delta}$ with $\delta_{e_1} = d_1$, $\delta_{e_2} = d_2$ and $\delta_{e_3} = d_3$,

$$e^{-\varepsilon|\mathcal{M}_2|/(2|\mathcal{M}|)}\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}) \le \zeta \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) \le e^{\varepsilon|\mathcal{M}_2|/(2|\mathcal{M}|)}\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}).$$

Note that our estimate for z_0 gives an estimate for $\zeta = z_0/\sqrt{R/S_1S_2S_3}$ with accuracy parameter at most $\varepsilon|\mathcal{M}_2|/(2|\mathcal{M}|)$. ($R$, S_1, S_2 and S_3 are byproducts of the computation of d_1, d_2 and d_3.)

Step 6a: Recall from Section 4.1 that $\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_3, \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{M}_1, \boldsymbol{\gamma}')$, where $\boldsymbol{\gamma}'$ is derived from $\boldsymbol{\gamma}_1$ by assigning $\gamma'_{p_1} = \delta_{e_1}$, $\gamma'_{p_2} = \delta_{e_2}$ and $\gamma'_{p_3} = \delta_{e_3}$.

Step 7a: Recursively estimate $\widetilde{Z}(\mathcal{M}_1; \boldsymbol{\gamma}')$ with accuracy parameter $\varepsilon(|\mathcal{M}| - |\mathcal{M}_2|)/|\mathcal{M}|$ and multiply it by the estimate for ζ from Step 5a. Return this value, which is an estimate of $\widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma})$.

FIGURE 2. The 3-sum case.

Step 4b: Recursively estimate $z_0 = \widetilde{Z}(\mathcal{M}_2 \backslash p; \boldsymbol{\gamma}_2)$ and $z_1 = \widetilde{Z}(\mathcal{M}_2/p; \boldsymbol{\gamma}_2)$ with accuracy parameter $(\varepsilon\varrho|\mathcal{M}_2|)/(2|\mathcal{M}|)$.

Step 5b: Using Lemma 13, compute d such that for every weight function $\boldsymbol{\delta}$ with $\delta_e = d$,

$$e^{-\varepsilon|\mathcal{M}_2|/(2|\mathcal{M}|)}\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}) \le \zeta \widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2; \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) \le e^{\varepsilon|\mathcal{M}_2|/(2|\mathcal{M}|)}\widetilde{Z}(\mathcal{M};\boldsymbol{\gamma}).$$

Note that our estimate for z_0 gives an estimate for $\zeta = 2(2+d)^{-1}z_0$ with accuracy parameter at most $(\varepsilon|\mathcal{M}_2|)/(2|\mathcal{M}|)$.

Step 6b: Recall from Section 4.1 that $\widetilde{Z}(\mathcal{M}_1 \triangle \mathcal{I}_2, \boldsymbol{\gamma}_1 \triangle \boldsymbol{\delta}) = \widetilde{Z}(\mathcal{M}_1, \boldsymbol{\gamma}')$, where $\boldsymbol{\gamma}'$ is derived from $\boldsymbol{\gamma}_1$ by assigning $\gamma'_p = d$.

Step 7b: Recursively estimate $\widetilde{Z}(\mathcal{M}_1; \boldsymbol{\gamma}')$ with accuracy parameter $\varepsilon(|\mathcal{M}| - |\mathcal{M}_2|)/|\mathcal{M}|$ and multiply it by the estimate for ζ from Step 5b. Return this value, which is an estimate of $\widetilde{Z}(\mathcal{M}, \boldsymbol{\gamma})$.

FIGURE 3. The 2-sum case.

The rest of this section shows that the running time is at most a polynomial in $|E(\mathcal{M})|$ and ε^{-1}.

Let $T_{\text{decomp}}(m) = O(m^{\alpha_{\text{decomp}}})$ be the time complexity of performing the Seymour decomposition of an m-element matroid (this is our initial splitting step), and let $T_{\text{base}}(m, \varepsilon) = O(m^{\alpha_{\text{base}}}\varepsilon^{-2})$ be the time complexity of estimating the Ising partition function of an m-edge graph. (From [8, Theorem 5], and the remark following it,

we may take $\alpha_{\rm base} = 15$.) Denote by $T(m,\varepsilon)$ the time-complexity of the algorithm of Figure 1. Recall that ϱ is a sufficiently small positive constant which does not depend upon $\mathcal{M}$ or ε and is described in Lemmas 12 and 13; we can take it to be $\varrho = 1/6000$. The recurrence governing $T(m,\varepsilon)$ is now presented, immediately followed by an explanation of its various components.

$$\begin{aligned} T(m,\varepsilon) \le T_{\rm decomp}(m) + \max\Big\{ & T_{\rm base}(m,\varepsilon), \\ & \max_{4\le k\le m/2}\big(T(m-k+3, \tfrac{\varepsilon(m-k-3)}{m}) + 4T(k, \tfrac{\varepsilon\varrho(k+3)}{4m})\big), \\ & \max_{2\le k\le m/2}\big(T(m-k+1, \tfrac{\varepsilon(m-k-1)}{m}) + 2T(k, \tfrac{\varepsilon\varrho(k+1)}{2m})\big), \\ & \max_{1\le k\le m/2}\big(T(m-k, \tfrac{\varepsilon(m-k)}{m}) + T(k, \tfrac{\varepsilon k}{m})\big)\Big\}. \end{aligned}$$

The four expressions within the outer maximisation correspond to the direct case, the 3-sum case, the 2-sum case, and the 1-sum case, respectively. The variable k is to be interpreted as the number of ground set elements in $\mathcal{M}$ that come from $\mathcal{M}_2$ (and hence $m-k$ is the number that come from $\mathcal{M}_1$). Thus, in the case of a 3-sum, for example, $|E(\mathcal{M}_1)| = m-k+3$ and $|E(\mathcal{M}_2)| = k+3$. Note that Step 2 of the algorithm ensures $k \le m/2$. The lower bounds on k come from the corresponding lower bounds on the size of matroids occurring in 3-sums, 2-sums and 1-sums.

We will demonstrate that $T(m,\varepsilon) = O(m^\alpha\varepsilon^{-2})$, where $\alpha = \max\{\alpha_{\rm base}, \alpha_{\rm decomp}+1, 43\}$. Specifically, we will show, by induction on m, that $T(m,\varepsilon) \le Cm^\alpha\varepsilon^{-2}$, for some constant C and all sufficiently large m. In the analysis that follows we do not attempt to obtain the best possible exponent α for the running time, instead preferring to simplify the analysis as much as possible. It would certainly be possible to reduce the constant 43 appearing in the formula for the exponent, but there seems little point in doing so, as the best existing value for $\alpha_{\rm base}$ is already too large to make the algorithm feasible in practice.

Note that by choosing C sufficiently large in the time-bound $T(m,\varepsilon) \le Cm^\alpha\varepsilon^{-2}$, we may ensure that this bound on time-complexity bound holds for any m from a finite initial segment of the positive integers, specifically for $m \in \{1, 2, \dots, 19\}$. For the inductive step, substitute $T(m,\varepsilon) \le Cm^\alpha\varepsilon^{-2}$ into the right-hand side of the above recurrence. We need to verify

$$Cm^\alpha\varepsilon^{-2} \ge T_{\rm decomp}(m) + \max\Big\{T_{\rm base}(m,\varepsilon), \tag{30}$$

$$\max_{4\le k\le m/2}\big(C(m-k+3)^\alpha\big(\tfrac{m}{\varepsilon(m-k-3)}\big)^2 + 4Ck^\alpha\big(\tfrac{4m}{\varepsilon\varrho(k+3)}\big)^2\big), \tag{31}$$

$$\max_{2\le k\le m/2}\big(C(m-k+1)^\alpha\big(\tfrac{m}{\varepsilon(m-k-1)}\big)^2 + 2Ck^\alpha\big(\tfrac{2m}{\varepsilon\varrho(k+1)}\big)^2\big) \tag{32}$$

$$\max_{1\le k\le m/2}\big(C(m-k)^\alpha\big(\tfrac{m}{\varepsilon(m-k)}\big)^2 + Ck^\alpha\big(\tfrac{m}{\varepsilon k}\big)^2\big)\Big\}, \tag{33}$$

for $m \geq 20$, as this will imply $T(m,\varepsilon) \leq Cm^\alpha\varepsilon^{-2}$, completing the induction step. Effectively, there are four independent inequalities to verify, numbered (30)–(33). The first inequality is immediate (for a sufficiently large constant C), since we are assuming $\alpha \geq \max\{\alpha_{\mathrm{decomp}}, \alpha_{\mathrm{base}}\}$.

The second inequality, namely (31), requires some work, but the remaining two will then follow easily. First, note the following simple estimate:

$$\frac{m-k+3}{m-k-3} \leq \frac{m/2+3}{m/2-3} = 1 + \frac{12}{m-6} \leq 1 + \frac{20}{m}, \tag{34}$$

where we have assumed that $k \leq m/2$ and $m \geq 20$. The inequality we wish to establish is equivalent, under rearrangement, to

$$Cm^\alpha\varepsilon^{-2} - C(m-k+3)^\alpha\big(\tfrac{m}{\varepsilon(m-k-3)}\big)^2 - 4Ck^\alpha\big(\tfrac{4m}{\varepsilon\varrho(k+3)}\big)^2 \geq T_{\mathrm{decomp}}(m),$$

for all k with $4 \leq k \leq m/2$. Noting (34), it is enough to show

$$Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-k+3)^{\alpha-2}(1+20/m)^2 - 64k^{\alpha-2}\varrho^{-2}\big] \geq T_{\mathrm{decomp}}(m), \tag{35}$$

for all k with $4 \leq k \leq m/2$. Regarding the left-hand side of (35) as a continuous function of a real variable k, and taking the second derivative with respect to k, we that the left-hand side is a concave function of k. It is enough, then, to check that (35) holds at $k = 4$ and $k = m/2$.

When $k = 4$, the inequality follows from the following sequence of inequalities:

$$\begin{aligned}
&Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-1)^{\alpha-2}(1+20/m)^2 - 64 \times 4^{\alpha-2}\varrho^{-2}\big] && (36)\\
&\quad\geq Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-1)^{\alpha-2}(1+1/m)^{40} - 4^{\alpha+1}\varrho^{-2}\big]\\
&\quad= Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - m^{\alpha-2}(1-1/m)^{\alpha-2}(1+1/m)^{40} - 4^{\alpha+1}\varrho^{-2}\big]\\
&\quad\geq Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - m^{\alpha-2}(1-1/m) - 4^{\alpha+1}\varrho^{-2}\big] && (37)\\
&\quad= Cm^2\varepsilon^{-2}\big[m^{\alpha-3} - 4^{\alpha+1}\varrho^{-2}\big]\\
&\quad\geq \tfrac{1}{2}Cm^{\alpha-1}\varepsilon^{-2} && (38)\\
&\quad\geq T_{\mathrm{decomp}}(m), && (39)
\end{aligned}$$

where inequality (37) is a consequence of $\alpha \geq 43$, inequality (39) of $\alpha \geq \alpha_{\mathrm{decomp}}+1$, and inequality (38) comes from a comparison of the two terms, noting $\alpha \geq 43$, $\varrho \geq 1/6000$ and $m \geq 20$.

When $k = m/2$, inequality (35) is established as follows:

$$\begin{aligned}
&Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m/2+3)^{\alpha-2}(1+20/m)^2 - (m/2)^{\alpha-2}(8/\varrho)^2\big] && (40)\\
&\quad= Cm^\alpha\varepsilon^{-2}\big[1 - (\tfrac{1}{2}+3/m)^{\alpha-2}(1+20/m)^2 - (1/2)^{\alpha-2}(8/\varrho)^2\big]\\
&\quad\geq Cm^\alpha\varepsilon^{-2}\big[1 - 10^{-7} - 10^{-2}\big] && (41)\\
&\quad\geq \tfrac{1}{2}Cm^\alpha\varepsilon^{-2}\\
&\quad\geq T_{\mathrm{decomp}}(m), && (42)
\end{aligned}$$

where (42) is a consequence of $\alpha \geq \alpha_{\text{decomp}}$, and (41) of $\alpha \geq 43$, $\varrho \geq 1/6000$ and $m \geq 20$.

The above calculations may easily be adapted to cover inequalities (32) and (33) for the cases of 2-sum and 1-sum. Since estimate (34) applies as well to $(m-k+1)/(m-k-1)$, the analogue of (35) in the 2-sum case is

$$Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-k+1)^{\alpha-2}(1+20/m)^2 - 8k^{\alpha-2}\varrho^{-2}\big] \geq T_{\text{decomp}}(m).$$

Again, we need only verify this at the extreme values of k, namely $k=2$ and $k=m/2$. The specialisation to $k=2$,

$$Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-1)^{\alpha-2}(1+20/m)^2 - 2^{\alpha+1}\varrho^{-2}\big] \geq T_{\text{decomp}}(m),$$

can be seen by comparison with (36), and the one for $k=m/2$,

$$Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m/2+1)^{\alpha-2}(1+20/m)^2 - 8(m/2)^{\alpha-2}\varrho^{-2}\big] \geq T_{\text{decomp}}(m).$$

by comparison with (40).

Finally, the analogue of (35) in the 1-sum case is

$$Cm^2\varepsilon^{-2}\big[m^{\alpha-2} - (m-k)^{\alpha-2} - k^{\alpha-2}\big] \geq T_{\text{decomp}}(m),$$

which again needs to be verified at the extreme values $k=1$ and $k=m/2$. As before, these can be seen by comparison with (36) and (40).

References

[1] Artur Andrzejak. Splitting formulas for Tutte polynomials. *J. Combin. Theory Ser. B*, 70(2):346–366, 1997.

[2] Béla Bollobás. *Combinatorics*. Cambridge University Press, Cambridge, 1986. Set systems, hypergraphs, families of vectors and combinatorial probability.

[3] Leslie Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic potts model. In Samson Abramsky, Cyril Gavoille, Claude Kirchner, Friedhelm Meyer auf der Heide, and Paul Spirakis, editors, *Automata, Languages and Programming*, volume 6198 of *Lecture Notes in Computer Science*, pages 396–407. Springer Berlin / Heidelberg, 2010.

[4] Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial. *Inform. and Comput.*, 206(7):908–929, 2008.

[5] Leslie Ann Goldberg and Mark Jerrum. Approximating the Tutte polynomial of a binary matroid and other related combinatorial polynomials. arXiv:1006.5234v1 [cs.CC], 2010.

[6] Alexander Golynski and Joseph D. Horton. A polynomial time algorithm to find the minimum cycle basis of a regular matroid. In *Algorithm theory—SWAT 2002*, volume 2368 of *Lecture Notes in Comput. Sci.*, pages 200–209. Springer, Berlin, 2002.

[7] F. Jaeger, D. L. Vertigan, and D. J. A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Math. Proc. Cambridge Philos. Soc.*, 108(1):35–53, 1990.

[8] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993.

[9] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoret. Comput. Sci.*, 43(2-3):169–188, 1986.

[10] James G. Oxley. *Matroid theory*. Oxford Science Publications. The Clarendon Press Oxford University Press, New York, 1992.

[11] R. B. Potts. Some generalized order-disorder transformations. *Proc. Cambridge Philos. Soc.*, 48:106–109, 1952.
[12] P. D. Seymour. Decomposition of regular matroids. *J. Combin. Theory Ser. B*, 28(3):305–359, 1980.
[13] Alan D. Sokal. The multivariate Tutte polynomial (alias Potts model) for graphs and matroids. In *Surveys in combinatorics 2005*, volume 327 of *London Math. Soc. Lecture Note Ser.*, pages 173–226. Cambridge Univ. Press, Cambridge, 2005.
[14] K. Truemper. *Matroid decomposition.* Academic Press Inc., Boston, MA, 1992.
[15] Klaus Truemper. A decomposition theory for matroids. V. Testing of matrix total unimodularity. *J. Combin. Theory Ser. B*, 49(2):241–281, 1990.

6. Appendix

6.1. **Part one.** The first part of the appendix contains some additional calculations needed in the proofs of Lemmas 4 and 5 to provide lower bounds on the quantity $c(A,B)$. We believe that several people have already done this calculation. Essentially, it shows that the two special cases of delta-sum that we consider can be constructed via a "generalized parallel connection", which is a special amalgam of $\mathcal{M}_1$ and $\mathcal{M}_2$ called a "proper" amalgam. This fact is well-known. See, for example [10, Section 12.4] or [1]. We couldn't find the calculation written down anywhere, and our result relies on it, so we include the calculation here. We use some ideas from [12, Section 2].

6.1.1. *Lemma 4:* $N = \mathcal{N}_1$. Recall that $c(A,B) = r_{\mathcal{M}}(A\cup B) - r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B)$. Our goal is to show

$$c(A,B) \geq \begin{cases} -1 & \text{iff } e_1(A,\{p\}) = 0 \text{ and } e_2(B,\{p\}) = 0; \\ 0 & \text{otherwise.} \end{cases}$$

Our basic strategy will be to identify sets $A'' \subseteq A$ and $B'' \subseteq B$ such that $r_{\mathcal{M}_1}(A'') = |A''|$ and $r_{\mathcal{M}_2}(B'') = |B''|$ and

$$e_1(A'',\{p\}) + e_2(B'',\{p\}) > 0 \tag{43}$$

Equation (43) implies that $r_{\mathcal{M}}(A''\cup B'') = |A''\cup B''|$. (Otherwise, $A''\cup B''$ contains a non-trivial cycle C. By the definition of delta-sum, it is of the form $C = C_1 \oplus C_2$ with $C_1 \in \mathcal{C}(\mathcal{M}_1)$ and $C_2 \in \mathcal{C}(\mathcal{M}_2)$. Since A'' is independent in $\mathcal{M}_1$ and B'' is independent in $\mathcal{M}_2$, neither C_1 nor C_2 is trivial. Now C_1 is a subset of $A''\cup\{p\}$, so $A''\cup\{p\}$ is dependent in $\mathcal{M}_1$, but A'' is independent in $\mathcal{M}_1$, so $e_1(A'',p) = 0$. Similarly, $e_2(B'',p) = 0$, which is a contradiction.) Then

$$\begin{aligned} c(A,B) &= r_{\mathcal{M}}(A\cup B) - r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B) \\ &\geq |A''| + |B''| - r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B). \end{aligned} \tag{44}$$

To implement the strategy, we first choose $B'' \subseteq B$ with $r_{\mathcal{M}_2}(B) = r_{\mathcal{M}_2}(B'') = |B''|$. Choose $A' \subseteq A$ with $r_{\mathcal{M}_1}(A) = r_{\mathcal{M}_1}(A') = |A'|$. If $e_1(A',\{p\}) + e_2(B'',\{p\}) > 0$ (which is certainly true as long as $e_1(A,\{p\}) + e_2(B,\{p\}) > 0$) then we take $A'' = A'$ so (44) gives $c(A,B) \geq 0$.

Otherwise, $A' \cup \{p\}$ contains a circuit. Let $e \neq p$ be a member of this circuit and take $A'' = A' - e$. Then $r_{\mathcal{M}_1}(A'') = |A''| = |A'| - 1$ and $r_{\mathcal{M}_1}(A'' \cup \{p\}) = |A'|$ so $e_1(A'', \{p\}) > 0$. Equation (44) gives Then $c(A, B) \geq -1$.

6.1.2. *Lemma 5: $N = \mathcal{N}_3$.* Our goal is to show

$$C \leq \begin{pmatrix} q^2 & q & q & q & 1 \\ q & q & 1 & 1 & 1 \\ q & 1 & q & 1 & 1 \\ q & 1 & 1 & q & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

where $C_{j,k} = q^{-c(A,B)}$, for $0 \leq j, k \leq 4$, where $A \subseteq E_1$ and $B \subseteq E_2$ are any sets satisfying $\varphi^1_j(A)$ and $\varphi^2_k(B)$.

The basic strategy is similar to the one used in Appendix 6.1.1. We wish to identify sets $A'' \subseteq A$ and $B'' \subseteq B$ such that $r_{\mathcal{M}_1}(A'') = |A''|$ and $r_{\mathcal{M}_2}(B'') = |B''|$ and, for every $p_j \in T$,

$$e_1(A'', \{p_j\}) + e_2(B'', \{p_j\}) > 0. \tag{45}$$

Equation (45) implies that $r_{\mathcal{M}}(A'' \cup B'') = |A'' \cup B''|$. The reasoning is a little more complicated than in Appendix 6.1.1, but it is similar. If $A'' \cup B''$ is not independent in $\mathcal{M}$ then, as before, $A'' \cup B''$ contains a non-trivial cycle $C = C_1 \oplus C_2$ where the non-trivial cycle $C_1 \in \mathcal{C}(\mathcal{M}_1)$ is a subset of $A'' \cup T$ and the non-trivial cycle $C_2 \in \mathcal{C}(\mathcal{M}_2)$ is a subset of $B'' \cup T$. Note that $C = (C_1 \oplus T) \oplus (C_2 \oplus T)$ so we can assume, without loss of generality, that $|C_1 \cap T| \leq 1$ (otherwise, replace C_1 with $C_1 \oplus T$ and replace C_2 with $C_2 \oplus T$). Also, $C_1 \cap T$ is nonempty. Otherwise, $C_1 \subseteq A''$, contradicting the independence of A'' in $\mathcal{M}_1$. So suppose $C_1 \cap T = \{p_j\}$. We can deduce that $C_2 \cap T = \{p_j\}$, since C avoids T. Now, as in Appendix 6.1.1, we get a contradiction. Since $C_1 \subseteq A'' \cup \{p_j\}$, but A'' is independent in $\mathcal{M}_1$, we get $e_1(A'', p_j) = 0$. Similarly, $e_2(B'', p_j) = 0$, which is a contradiction. As in Appendix 6.1.1, implementing the basic strategy allows us to conclude

$$c(A, B) \geq |A''| + |B''| - r_{\mathcal{M}_1}(A) - r_{\mathcal{M}_2}(B). \tag{46}$$

To implement the strategy, we choose $B' \subseteq B$ with $r_{\mathcal{M}_2}(B) = r_{\mathcal{M}_2}(B') = |B'|$ and we choose $A' \subseteq A$ with $r_{\mathcal{M}_1}(A) = r_{\mathcal{M}_1}(A') = |A'|$.

If, for all $p_j \in T$, $e_1(A', \{p_j\}) + e_2(B', \{p_j\}) > 0$, (which is certainly true as long as, for all $p_j \in T$, $e_1(A, \{p_j\}) + e_2(B, \{p_j\}) > 0$) then we take $A'' = A'$ and $B'' = B'$ so (46) gives $c(A, B) \geq 0$. This accounts for all of the "1" entries in the matrix above.

Suppose, instead, that $e_1(A', \{p_1\}) + e_2(B', \{p_1\}) = 0$ but $e_1(A', \{p_2\}) + e_2(B', \{p_2\}) > 0$ and $e_1(A', \{p_3\}) + e_2(B', \{p_3\}) > 0$. (This case arises for all cases with $e_1(A, \{p_2\}) + e_2(B, \{p_2\}) > 0$ and $e_1(A, \{p_3\}) + e_2(B, \{p_3\}) > 0$ that we have not already covered.) In this case, $A' \cup \{p_1\}$ contains a circuit. Let $e \neq p_1$ be a member of this circuit and take $A'' = A' - e$. Then $r_{\mathcal{M}_1}(A'') = |A''| = |A'| - 1$ and

$r_{\mathcal{M}_1}(A'' \cup \{p_1\}) = |A'|$ so $e_1(A'', \{p_1\}) > 0$. Taking $B'' = B'$, Equation (46) gives Then $c(A, B) \geq -1$. By symmetry (swapping p_1 with either p_2 or p_3), this accounts for all of the "q" entries in the matrix above.

For the final case, suppose that, for all $p_j \in T$, $e_1(A', \{p_j\}) + e_2(B', \{p_j\}) = 0$. Define A'' as above so that $e_1(A'', \{p_1\}) > 0$. As we note just after enumerating the possibilities P0, ..., P4 in the proof of Lemma 5, it is not possible that exactly one of $e_1(A'', \{p_1\})$, $e_1(A'', \{p_2\})$, and $e_1(A'', \{p_3\})$ is 1, so assume without loss of generality that $e_1(A'', \{p_2\}) > 0$. Now $B' \cup \{p_3\}$ contains a circuit. Let $f \neq p_3$ be a member of this circuit and take $B'' = B' - f$ so $r_{\mathcal{M}_2}(B'') = |B''| = |B'| - 1$ and $e_2(B'', \{p_3\}) > 0$. Equation (46) now gives $c(A, B) \geq -2$, which accounts for the q^2 entry in the matrix.

6.2. **Part two.** The second part of the appendix contains some technical lemmas needed in the proof of Lemma 12. The lemmas are about approximation, and they don't add any intuition to the paper.

We start with a well-known fact that we will use in the proof of both lemmas. (This follows directly from the series expansion of e.)

Observation 14. *If* $0 < \varepsilon < 1$ *then* $1 + \varepsilon \leq e^{\varepsilon} \leq 1 + 2\varepsilon$.

Our first lemma refers to the matrix D defined just before Lemma 5.

Lemma 15. *Suppose that* z*,* s *and* r *are column vectors in* $\mathbb{R}^4$ *with positive entries satisfying* $1 \leq z_i/z_0 \leq 2$*,* $1 \leq s_i/s_0 \leq 2$*, and* $1 \leq r_i/r_0 \leq 2$ *for* $i \in \{1, 2, 3\}$*. Suppose that* $e^{-\varepsilon} s_i \leq r_i \leq e^{\varepsilon} s_i$ *for some* $0 < \varepsilon < 1$*. Then* $e^{-56\varepsilon} z^{\mathrm{T}} D s \leq z^{\mathrm{T}} D r \leq e^{56\varepsilon} z^{\mathrm{T}} D s$*.*

Proof. For any positive column vector v in $\mathbb{R}^4$ satisying $1 \leq v_i/v_0$,

$$z^{\mathrm{T}} D v = z_0 v_0 + (z_1 - z_0)(v_1 - v_0) + (z_2 - z_0)(v_2 - v_0) + (z_3 - z_0)(v_3 - v_0) \geq z_0 v_0.$$

Also, summing the absolute values of all of the monomials in $z^{\mathrm{T}} D v$, and using $z_i/z_0 \leq 2$ and $v_i/v_0 \leq 2$, we get $\sum_{i,j} |D_{i,j}| z_i v_j \leq 28 z_0 v_0$. Then

$$\begin{aligned} z^{\mathrm{T}} D s - z^{\mathrm{T}} D r &= \sum_{i,j} D_{i,j} z_i (s_j - r_j) \\ &\leq \sum_{i,j} |D_{i,j}| z_i (e^{\varepsilon} - 1) r_j \\ &\leq 2\varepsilon \sum_{i,j} |D_{i,j}| z_i r_j \\ &\leq 56 \varepsilon z_0 r_0 \\ &\leq 56 \varepsilon z^{\mathrm{T}} D r, \end{aligned}$$

so $z^{\mathrm{T}} D s \leq (1 + 56\varepsilon) z^{\mathrm{T}} D r \leq e^{56\varepsilon} z^{\mathrm{T}} D r$. □

Our second lemma refers to the equations in Lemma 10.

Lemma 16. *Suppose that $r_1 \le r_2 \le r_3$ satisfy the following equations.*

$$2 + r_1 - r_2 - r_3 > 0, \tag{47}$$

$$r_1 + r_2 + r_3 - r_2 r_3 - 2 \ge 0, \tag{48}$$

$$1 \le r_i \le 2, \text{ for } i \in \{1, 2, 3\}. \tag{49}$$

Suppose that we are given values $\tilde{s}_1 \le \tilde{s}_2 \le \tilde{s}_3$ satisfying $e^{-\chi} r_i \le \tilde{s}_i \le e^{\chi} r_i$ for $i \in \{1, 2, 3\}$, where χ is a sufficiently small positive constant. Using $\tilde{s}_1$, $\tilde{s}_2$ and $\tilde{s}_3$, we can compute values s_1, s_2 and s_3 satisfying $1 \le s_i \le 2$,

$$e^{-66\chi} s_i \le r_i \le e^{66\chi} s_i,$$

and Equations (25), (26) and (27).

Proof. Let $\delta = 4e\chi$. Note that

$$|\tilde{s}_i - r_i| \le \delta. \tag{50}$$

For example, using Observation 14, $r_i - \tilde{s}_i \le 2\chi \tilde{s}_i \le 2\chi e^{\chi} r_i$. Since $r_i \le 2$ (by Equation (49)) and $\chi \le 1$, $r_i - \tilde{s}_i \le \delta$. Similarly, $\tilde{s}_i - r_i \le \delta$.

In each of two cases, we will compute s_1, s_2, and s_3 so that

$$1 \le s_1 \le s_2 \le s_3 \le 2, \tag{51}$$

$$2 + s_1 - s_2 - s_3 > 0, \tag{52}$$

$$s_1 + s_2 + s_3 - s_2 s_3 - 2 \ge 0, \tag{53}$$

$$|s_i - r_i| \le 6\delta. \tag{54}$$

Equations (51), (52) and (53) imply Equations (25), (26) and (27). Also, Equation (54) implies $e^{-66\chi} s_i \le r_i \le e^{66\chi} s_i$ since $s_i \le r_i + 6\delta \le r_i(1 + 6\delta) \le r_i e^{6\delta} \le r_i e^{66\chi}$ and similarly $r_i \le s_i e^{66\chi}$.

Case 1: $\tilde{s}_2 - \tilde{s}_1 \le 5\delta$.

Take $s_1 = s_2 = \min(\max(1, \tilde{s}_2), 2 - \delta)$ and $s_3 = \min(\max(1, \tilde{s}_3), 2 - \delta)$. (51) follows easily from the definitions since $2 - \delta \ge 1$ and $\tilde{s}_2 \le \tilde{s}_3$. (52) and (53) follow from the facts that $s_1 = s_2$, $s_1 \ge 1$ and $s_3 < 2$ since, $s_1 + s_1 + s_3 - s_1 s_3 - 2 = (s_1 - 1)(2 - s_3)$. To establish (54) note that

$$s_1 \le \max(1, \tilde{s}_2) \le \max(1, \tilde{s}_1 + 5\delta) \le \max(1, r_1 + 6\delta) \le r_1 + 6\delta.$$

Also, since $\tilde{s}_2 \le r_2 + \delta \le 2 + \delta$,

$$s_1 \ge \tilde{s}_2 - 2\delta \ge \tilde{s}_1 - 2\delta \ge r_1 - 3\delta.$$

Similarly, $r_2 - 3\delta \le s_2 \le r_2 + \delta$ and $r_3 - 3\delta \le s_3 \le r_3 + \delta$.

Case 2: $\tilde{s}_2 - \tilde{s}_1 > 5\delta$.

Take $s_1 = \tilde{s}_1 + 4\delta$, $s_2 = \min(\tilde{s}_2, 2)$ and $s_3 = \min(\tilde{s}_3, 2)$. Equation (51) follows since $\tilde{s}_1 + 4\delta \leq \tilde{s}_2 - \delta \leq s_2$. (52) follows since

$$\begin{aligned} 2 + s_1 - s_2 - s_3 &\geq 2 + \tilde{s}_1 + 4\delta - \tilde{s}_2 - \tilde{s}_3 \\ &\geq 2 + (r_1 - \delta) + 4\delta - (r_2 + \delta) - (r_3 + \delta) \\ &\geq \delta + 2 + r_1 - r_2 - r_3 \\ &> \delta. \end{aligned}$$

We now show that (53) holds. Note that

$$s_1 + s_2 + s_3 - s_2 s_3 - 2 = s_1 - (s_2 - 1)(s_3 - 1) - 1,$$

and the this quantity is increasing as a function of s_1 and decreasing as a function of s_2 and as a function of s_3. Also,

$$\begin{aligned} &s_1 = \tilde{s}_1 + 4\delta \geq r_1 + 3\delta, \\ &s_2 = \min(\tilde{s}_2, 2) \leq \tilde{s}_2 \leq r_2 + \delta, \text{ and} \\ &s_3 \leq r_3 + \delta. \end{aligned}$$

So

$$\begin{aligned} s_1 + s_2 + s_3 - s_2 s_3 - 2 &\geq (r_1 + 3\delta) - (r_2 + \delta - 1)(r_3 + \delta - 1) - 1 \\ &= 3\delta - \delta(r_2 - 1) - \delta(r_3 - 1) - \delta^2 + (r_1 - (r_2 - 1)(r_3 - 1) - 1) \\ &\geq 3\delta - \delta(r_2 - 1) - \delta(r_3 - 1) - \delta^2, \end{aligned}$$

and this is at least 0, since $r_2 - 1 \leq 1$, $r_3 - 1 \leq 1$, and $\delta \leq 1$.

To establish (54), note that $s_1 = \tilde{s}_1 + 4\delta \leq r_1 + 5\delta$ and $r_1 \leq s_1$. Also, as noted above, $\tilde{s}_2 - \delta \leq s_2 \leq \tilde{s}_2$ and similarly $\tilde{s}_3 - \delta \leq s_3 \leq \tilde{s}_3$.

□

Department of Computer Science, University of Liverpool, Ashton Building, Liverpool L69 3BX, United Kingdom.

School of Mathematical Sciences, Queen Mary, University of London, Mile End Road, London E1 NS, United Kingdom.